THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Ribonuclease H2, RNA:DNA hybrids and innate immunity

Rachel Elizabeth Rigby

Thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh

2010

This thesis is composed of original research undertaken by myself, and where the work of others is included their contributions have been duly acknowledged. Some of the content of the Introduction of this thesis formed the basis of a review article of which I am first author.

Rachel Rigby
September 2010

Acknowledgements

The first big thank you has to go to Andrew Jackson, for being such a great supervisor, and for putting up with me for the past four years! And for all the help, ideas, encouragement and doughnuts provided throughout my PhD, and for being so very patient with my approach to deadlines! I have enjoyed my time in the lab immensely (paper-mania times included).

Of course, a massive thanks to all members of the Jackson lab, past and present, for all being so lovely, helpful and generally fantastic, especially at keeping me sane with lunches, pub trips and Wispa Golds during this last year. So, (name-checking in alphabetical order): Andrea, for making sure I always had whatever I needed in the lab, no matter how random, and mostly for always being there when I needed someone to talk/moan to and being so supportive for the whole four years you've had to sit next to me!; Anna, for thesis writing moral support and general (mostly unintentional!) entertainment; Bertrand, for patiently reminding me how to switch the fluorescent microscopes on every time I made my annual trip to the microscope room, for telling me all I will ever need to know about flies ('le Curly O') and unintentionally teaching me how to swear in French; Björn; for help optimising ELISAs, setting up MEFs, for providing material from the *Rnaseh2b* knock-in mice, commenting on my thesis and for always being so interested in hybrids and my experiments; Carol-Anne, my fellow late night and weekend lab buddy! Thanks for making evenings in the lab more entertaining and girly with shoes, handbags, make up and vampire chat! And for cracking me up with a new Carol-Anne-ism every once in a while; Ellie, for lots of help when I first started and for bequeathing me the Lab Faqs bible; Katy, for being a calming influence and for lending me Lil Pig to keep me company for a bit when I was writing up!; Louise-Kiwifruit, big thanks for all your help reading my thesis drafts (introducing me to the use of a space before numbers and units!), for advice regarding PCR optimisation (Queen of the PCR), and for teaching me about the divoone NZ foods (Anzacs, lamb burgers) and the ones to avoid (Pineapple Lumps!); Maggie, thanks for thesis writing tips and 'you can do it!' encouragement during the writing time; Martin, thanks for massive amounts of help with all things nucleic acid related: for protocols, advice, thesis draft reading, answering many questions by email and not forgetting the supplying of Stroopwafles/gardening advice; and finally to Rachel W, for providing recombinant RNase H2 and RNase assay protocols.

The next big thank you is to Andrew MacDonald at the University of Edinburgh for help with all things DC-related, including ideas, protocols, reagents, discussions, letting me do the FL-DC experiments in his lab and for being so enthusiastic about my hybrid experiments. Thanks also to Andrew's lab for all their help, especially Rachel Lundie for the FL-DC culture protocol and help getting my experiments up and running, and to Lauren Webb for organising me and helping me every time I forgot how to use the FACS machine or couldn't figure out the antibody colours! And for the occasional tap-dancing lesson during spins.

Thanks to David Gray at the University of Edinburgh for supplying the MyD88/Trif knock-out mice, and at the CRUK London Research Institute to Caetano Reis e Sousa for providing IPS-1 knock out bone marrow and wild type controls, to Neil Rogers for harvesting and sending the bone marrow and Delphine Goubau for the IVT RNA control for the IPS-1 experiments.

At the HGU, a big thanks to everyone in the E2 lab for all the help, reagents and friendly (!) football banter over the years, and for putting up with my ELISA-related noise pollution and slight obsession over the fridge magnets. Especially big thanks to Anne for all her help with cell culture, finding equipment, ordering things and for making the most amazing birthday cakes.

Thanks also to Julia Dorin for immunology discussions and lots of ideas for experiments over the years, and to Fiona Semple for sharing reagents and protocols. Thanks to Graeme Grimes for performing the initial analyses of microarray data and Kirstie Lawson for dissecting E8.5 and E9.5 embryos, John Wiltshire for help with setting up the FPLC protocols and hydroxyapatite column experiments, Lizzie Fryer for flow cytometry assistance, Stewart MacKay for running samples on the BioAnalyzer and help setting up a giant PCR and to Craig Nicol for many lessons on using Illustrator and InDesign.

Thanks to all the HGU students past and present for some great parties (piratical and non-piratical!), the fun nights in Hectors and some slightly more dubious nights in Garibaldis (especially our first MRes night out there which for some reason we all remember with great fondness!). Especially thanks to Hannah for all the free tickets to various things she's given me over the years, thesis writing tips and cracking food, and to James, for fixing my internet at every flat I have lived in!

Thanks also to my friends; the Edinburgh ladies Jen and Catherine for being so lovely and supportive, especially during this last year. Thanks for great Friday nights out dancing, making me amazing dinners and just generally being there. And of course thanks to the Liverpool ladies; Hel, Gel, Gem and Lucifer for all your trips up to Edinburgh to visit me and your texts of encouragement: promise I will see more of you all now it's written!

Lastly, the biggest thank you is to my Mum and Dad. Thank you so much for all your encouragement, support (including financial!) and help during my almost ten (!) years as a student. Thanks for always being interested in what I'm doing (even when you weren't completely sure what I was going on about!) and always being there for me to moan to after I've had a bad day and making me feel better. I couldn't have done it without you.

Abstract

The activation of the innate immune system is the first line of host defence against infection. Nucleic acids can potently stimulate this response and trigger a series of signalling cascades leading to cytokine production and the establishment of an inflammatory state. Mutations in genes encoding nucleases have been identified in patients with autoimmune diseases, including Aicardi-Goutières syndrome (AGS). This rare childhood inflammatory disorder is characterised by the presence of high levels of the antiviral cytokine interferon-α in the cerebrospinal fluid and blood, which is thought to be produced as a consequence of the activation of the innate immunity by unprocessed self-nucleic acids. This thesis therefore aimed to define the role of one of the AGS nucleases, the Ribonuclease H2 (RNase H2) complex, in innate immunity, and to establish if nucleic acid substrates of this enzyme were able to induce type I interferon production *in vitro*.

The AGS nucleases may function as components of the innate immune response to nucleic acids. Consistent with this hypothesis, RNase H2 was constitutively expressed in immune cells, however, its expression was not upregulated in response to type I interferons. RNase H2-deficient cells responded normally to a range of nucleic acid PAMPs, which implied that a role for RNase H2 as a negative regulator of the immune response was unlikely, in contrast to the reported cellular functions of two other AGS proteins, TREX1 and SAMHD1. Therefore, no clear evidence was found for the direct involvement of RNase H2 in the innate immune response to nucleic acids.

An alternative model for the pathogenesis of disease hypothesises that decreased RNase H2 activity within the cell results in an accumulation of RNA:DNA hybrids. To investigate the immunostimulatory potential of such substrates, RNA:DNA hybrids with different physiochemical properties were designed and synthesised. Methods to purify the hybrids from other contaminating nucleic acid species were established and their capacity as activators of the innate immune response tested

using a range of *in vitro* cellular systems. A GU-rich 60 bp RNA:DNA hybrid was shown to be an effective activator of a pro-inflammatory cytokine response exclusively in Flt3-L bone marrow cultures. This response was completely dependent on signalling involving MyD88 and/or Trif, however the specific receptor involved remains to be determined. Reduced cellular RNase H2 activity did not affect the ability of Flt3-L cultures to mount a cytokine response against the RNA:DNA hybrid. These *in vitro* studies suggested that RNA:DNA hybrids may be a novel nucleic acid PAMP. Taken together, the data in this thesis suggest that the cellular function of RNase H2 is in the suppression of substrate formation rather than as a component of the immune response pathways. Future studies to identify endogenous immunostimulatory RNA:DNA hybrids and the signalling pathways activated by them should provide a detailed understanding of the molecular mechanisms involved in the pathogenesis of AGS and related autoimmune diseases.

Table of Contents

List of Figures

List of tables

Abbreviations

AGS	Aicardi-Goutières syndrome
APS	ammonium persulphate
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BPB	bromophenol blue
BSA	bovine serum albumin
cDC	conventional/myeloid dendritic cell
cDNA	complementary DNA
CNS	central nervous system
CO_2	carbon dioxide
CSF	cerebrospinal fluid
Ct	cycle threshold
CT	computerised tomography
DABCYL	4-[[4-(Dimethylamino)phenyl]azo]benzoic acid
DAMP	danger-associated molecular pattern
DAPI	4',6-diamidino-2-phenylindole
DC	dendritic cell
^{o}C	degrees centigrade
DEPC	diethyl dicarbonate
ddH_2O	distilled deionised water
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
ds	double-stranded
DSB	double-strand break
dsNA	double-stranded nucleic acid
E	embryonic day
EDTA	ethylenediaminetetraacetic acid
EMEM	Eagle’s minimum essential medium
ER	endoplasmic reticulum
EtOH	ethanol
FAM	fluorescein amidite
FCS	foetal calf serum
FPLC	fast performance liquid chromatography
g	gram
GM-CSF	granulocyte macrophage-colony stimulating factor
GO	Gene Ontology
h	hour
HPLC	high performance liquid chromatography
IC	immune complex
IFN	interferon

Ig	immunoglobulin
ISD	interferon stimulatory DNA
IU	international unit
IVT	in vitro transcription
kb	kilobases
kDa	kilodaltons
LCL	lymphoblastoid cell line
LPS	lipopolysaccharide
M	molar
mM	millimolar
µM	micromolar
mAb	monoclonal antibody
M-CSF	macrophage-colony stimulating factor
mDC	myeloid dendritic cell
MEF	mouse embryonic fibroblast
mg	milligrams
$MgCl_2$	magnesium chloride
min	minutes
mJ	millijoule
mRNA	messenger ribonucleic acid
µl	microlitres
µm	micrometre
µM	micromolar
ml	millilitres
mM	millimolar
MOPS	3-(N-morpholino)propanesulfonic acid
NAC	no amplification control
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
ng	nanograms
nm	nanometre
nmol	nanomole
nt	nucleotide
NTC	no template control
OD	optical density
ODN	oligodeoxyribonucleotide
ORN	oligoribonucleotide
PAMP	pathogen associated molecular pattern
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PD	phosphodiester
pDC	plasmacytoid dendritic cell
PFA	paraformaldehyde
pmol	picomole
PRR	pattern recognition receptor
PS	phosphorothioate
R:D	RNA:DNA

RNA	ribonucleic acid
RNA:DNA	ribonucleic acid:deoxyribonucleic acid hybrid
RNase	ribonuclease
rNTP	ribonucleoside 5'-triphosphate
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute
rRNA	ribosomal RNA
RT	reverse transcriptase
qRT-PCR	quantitative real time polymerase chain reaction
s	second
s.d.	standard deviation
SDS	sodium dodecyl sulfate
s.e.m.	standard error of the mean
ss	single-stranded
ssNA	single-stranded nucleic acid
SLE	systemic lupus erythematosus
TBE	Tris/Borate/EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
TLR	Toll-like receptor
Tm	melting temperature
Tris	tris(hydroxymethyl)aminomethane
U	units
UTR	untranslated region
UV	ultraviolet
V	volt
v/v	volume per volume
W	Watt
w/v	weight per volume
XC	xylene cyanol FF
x g	x gravity

Chapter 1. Introduction

Enzymatic cleavage of most ribonucleic acid (RNA) molecules is necessary for producing their mature functional forms or conversely, for their degradation. These cleavage reactions are performed by cellular ribonucleases (RNases), essential components of all living organisms. Multiple classes of RNases acting upon a diverse range of substrates have been characterised extensively in *Escherichia coli* (*E. coli*) (reviewed by Nicholson, 1999). RNases can be broadly classified as either endoribonucleases, cleaving internal phosphodiester bonds in the backbone of the RNA molecule, or as exoribonucleases, which remove single ribonucleotides from the ends of the molecule.

In eukaryotes, RNase activity is a vital component of many cellular processes. Members of the RNase A superfamily exhibit antimicrobial activity (reviewed by Dyer and Rosenberg, 2006), whilst RNase L functions as an effector of the innate antiviral response (Zhou *et al.*, 1993) (see Section 1.4.3). The RNase III-like family of endoribonucleases are essential components of the RNA interference (RNAi) pathways (Bernstein *et al.*, 2001; Wu *et al.*, 2000). Defective enzymatic activity can be causative of human disease, exemplified by the congenital auto-inflammatory disorder Aicardi-Goutières syndrome (AGS) which can be caused by mutations in the genes encoding the subunits of a member of the RNase H class of enzymes (Crow *et al.*, 2006b).

1.1. Ribonucleases H

Ribonucleases H (RNases H) are endoribonucleases which specifically hydrolyse the RNA strand of an RNA:DNA hybrid (Hausen and Stein, 1970; Miller *et al.*, 1973). They are found in all kingdoms of life. Eukaryotic RNases H can be classified into Type 1 and Type 2 RNase H families based on amino acid similarity, size and biochemical properties (Ohtani *et al.*, 1999a). The major biochemical difference

between the two families is in the choice of diavalent cation required to bind substrate and catalyse nucleotidyl transfer: Type 2 enzymes can substitute Mn^{2+} as a cofactor in place of Mg^{2+} but Type 1 enzymes cannot and may be inhibited by the presence of Mn^{2+} (Eder and Walder, 1991; Nowotny *et al.*, 2005).

1.1.1. RNase H substrates

Both classes of RNase H enzymes are highly specific for RNA:DNA hybrids and hydrolyse the RNA strand in an endonucleolytic manner, however, Type 1 enzymes require at least four consecutive ribonucleoside 5′ triphosphates (rNTPs) for activity (Hogrefe *et al.*, 1990). In contrast, Type 2 RNases H are able to excise a single rNTP embedded in a DNA duplex, always hydrolysing the 5′ phosphodiester of the rNTP (Eder and Walder, 1991) (Figure 1.1). Additionally, it has been demonstrated that RNase H1 and RNase H2 exhibit different cleavage patterns on the same RNA:DNA hybrid substrates (Pileur *et al.*, 2000).

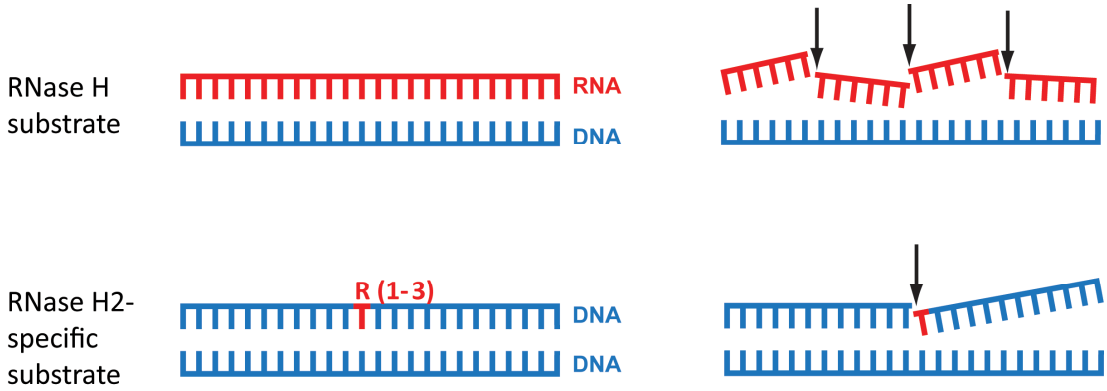


Figure 1.1. Differences in substrate choice for the two types of RNase H enzyme. Schematic representation of RNase H substrates. Both Type 1 and Type 2 enzymes are able to bind to and cleave RNA:DNA hybrids with a minimum of four consecutive ribonucleotides. In addition, Type 2 enzymes are able to recognise and cleave a single ribonucleotide incorporated into a DNA duplex. Sites of cleavage of RNase H2 are indicated with black arrows.

1.1.2. The RNA:DNA hybrid-cleaving activity of other proteins

The ubiquitous presence of RNase H enzymes in different organisms suggests this activity plays an important role in nucleic acid metabolism (Frank *et al.*, 1998a; Frank *et al.*, 1998b). RNase H activity is found in retroviruses as a crucial component of the reverse transcriptase (RT) that converts the ssRNA retroviral genome into dsDNA (Molling *et al.*, 1971; Wohrl and Molling, 1990). Related RNase H-like domains have been found in a number of other proteins, including the PIWI fold of Argonaute-2 (the protein component of the RNA-induced silencing complex (RISC)) (Parker *et al.*, 2004; Song *et al.*, 2004), and in the retroviral integrase and transposase enzymes (Rice and Baker, 2001).

1.1.3. Type 1 RNase H enzymes

Type 1 RNase H enzymes are present in all prokaryotes and eukaryotes, but are absent in most archaea (Ohtani *et al.*, 1999b; reviewed by Tadokoro and Kanaya, 2009). Human RNase H1 is a monomeric protein consisting of 286 amino acids with a molecular weight (MW) of approximately 32 kilodaltons (kDa) (Figure 1.2). All eukaryotic RNase H1 proteins contain a highly conserved region termed the hybrid binding domain (HBD). This nucleic acid binding domain is capable of binding dsRNA but shows a 25-fold preference for binding RNA:DNA hybrids, independent of their sequence (Cerritelli and Crouch, 1998; Nowotny *et al.*, 2008). Upon substrate binding, murine RNase H1 is induced to form a homodimeric enzyme (Gaidamakov *et al.*, 2005). The HBD contributes to the specific activity of RNase H1 by increasing the affinity of the enzyme for RNA:DNA hybrid substrates and so enhances the processivity of the enzyme (Cerritelli and Crouch, 2009; Gaidamakov *et al.*, 2005; Nowotny *et al.*, 2008).

The enzymatic C-termini of eukaryotic RNase H1 enzymes consist of approximately 150 highly conserved amino acids: the RNase H domain. The catalytic core contains four absolutely conserved carboxylic acid residues at the active site, which are essential for hydrolysis of substrates (Figure 1.2) (Katayanagi *et al.*, 1990; Wu *et al.*, 2001). This is highly similar to the catalytic site of prokaryotic RNase H1 enzymes

(Tadokoro and Kanaya, 2009). RNA:DNA hybrid recognition has been shown to be dependent on the identification of both the RNA strand (by the presence of the 2′-hydroxyl group (2′-OH)) and the DNA strand (by the absence of 2′-OH and/or the B-form conformation of the strand) (as discussed in Chapter 4) (Cerritelli and Crouch, 2009; Nowotny *et al.*, 2005; Nowotny *et al.*, 2007).

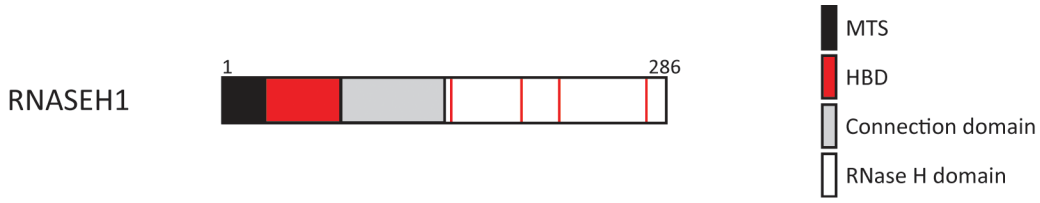


Figure 1.2. The human RNASEH1 protein
Schematic diagram showing human RNase H1. The 286 amino acid polypeptide contains a mitochondrial targeting sequence (MTS) at the extreme N-terminus, followed by a 50 amino acid hybrid binding domain (HBD). A 64 amino acid connection domain separates the HBD from the 150 amino acid RNase H domain. The locations of the four residues which form the active site are indicated with red lines at positions 145, 186, 210 and 274 amino acids (Wu *et al.*, 2001). The first three residues are essential for catalytic activity (Nowotny *et al.*, 2007; Wu *et al.*, 2001).

The N- and C-termini are separated by a connection domain, the length and sequence of which is variable between species. The primary role of this domain is to convey flexibility to the protein, so that the HBD and catalytic domain can be correctly positioned on the nucleic acid (Nowotny *et al.*, 2008), but it may also function in the interaction of RNase H1 with other proteins (Cerritelli and Crouch, 2009).

In higher eukaryotes, RNase H1 proteins contain a mitochondrial targeting sequence (MTS) located upstream of the HBD at the 5′-end which is coded for by an alternative start codon. It has been proven that this is capable of localising the enzyme to the mitochondria *in vitro* (Cerritelli *et al.*, 2003; Pileur *et al.*, 2000). RNase H1 has been shown to have an essential role during development for the generation of mitochondrial DNA (mtDNA) in mice (Cerritelli *et al.*, 2003) but not

in yeast which replicate mtDNA in an RNase H1-independent manner (Arudchandran *et al.*, 2000), discussed further in Section 1.1.5.

1.1.4. Type 2 RNase H enzymes

Type 2 RNase H activity is present in all organisms (Ohtani *et al.*, 1999b). The sequence of the catalytic site of the enzyme is highly conserved throughout evolution (Frank *et al.*, 1998b; Frank *et al.*, 1998c; Rydberg and Game, 2002). RNase H2 is ubiquitously expressed and constitutes the major source of cellular ribonuclease activity in eukaryotes (Arudchandran *et al.*, 2000; Eder and Walder, 1991; Frank *et al.*, 1998a; Frank *et al.*, 1998c).

1.1.4.1. The eukaryotic RNase H2 enzyme complex

Bacterial and archaeal RNase HII are active as monomeric proteins encoded for by the *rnhB* genes (Haruki *et al.*, 1998; Itaya, 1990), however eukaryotic RNase H2 requires additional protein subunits for functional enzymatic activity. The RNase H2 complex comprises three subunits; RNASEH2A, RNASEH2B and RNASEH2C (Crow *et al.*, 2006b; Frank *et al.*, 1998c; Jeong *et al.*, 2004) (Figure 1.3).

The catalytic RNASEH2A subunit has homologues in all eukaryotic species. The sequence and main fold of the protein share significant similarity to the prokaryotic and archaeal RNase HII proteins and use the same catalytic mechanism of four conserved residues coordinating the divalent cations at the active site (Chapados *et al.*, 2001; Frank *et al.*, 1998b; Jeong *et al.*, 2004; Lai *et al.*, 2000). The extreme N- and C-terminal regions of the eukaryotic RNASEH2A protein are absent in the bacterial and archaeal proteins (Jeong *et al.*, 2004), and significantly the C- terminus has been shown to interact with the B/C subunits (Shaban *et al.*, 2010) (Martin Reijns, MRC HGU, personal communication).

In contrast to RNASEH2A, the accessory proteins RNASEH2B and RNASEH2C share little sequence similarity with their orthologues and have no prokaryotic equivalents (Crow *et al.*, 2006b; Jeong *et al.*, 2004). However, they are absolutely

essential for RNase H2 activity in eukaryotes (Chon *et al.*, 2009; Crow *et al.*, 2006b; Jeong *et al.*, 2004). It has been proposed that they form a scaffold on which the A subunit can adopt the appropriate conformation for enzymatic activity (Chon *et al.*, 2009). It is also hypothesised that they may be essential for promoting the progressive action of the eukaryotic RNase H2 enzyme activity, as the monomeric prokaryotic RNase H2 is known to cleave RNA:DNA hybrids in a distributive manner (Chon *et al.*, 2009).

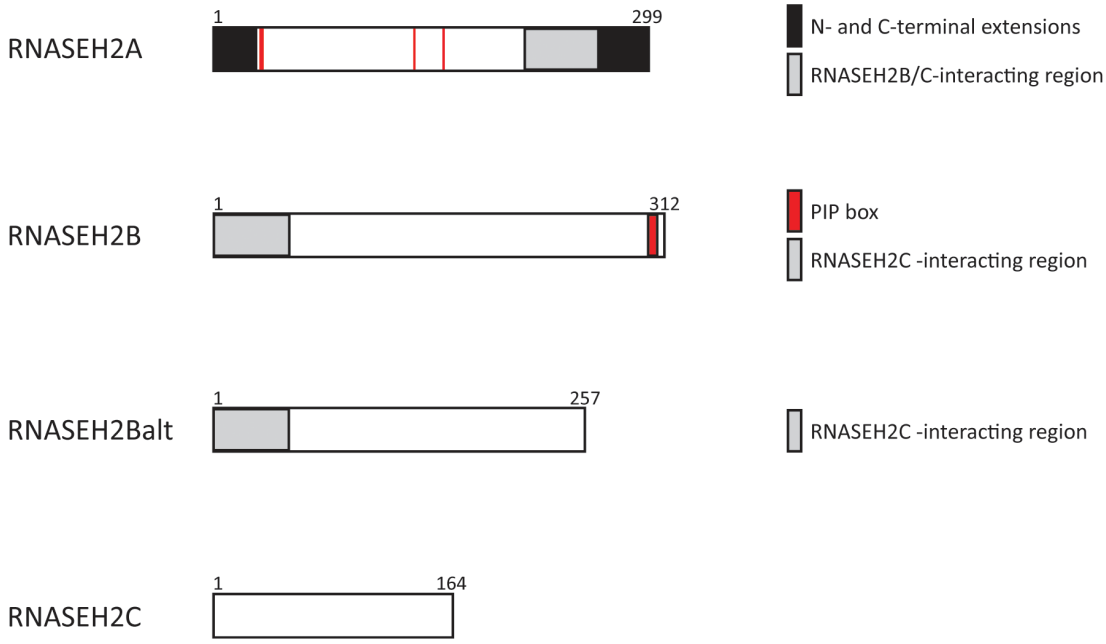


Figure 1.3. The human RNASEH2 proteins
Schematic representation of the three human RNase H2 proteins. The approximate locations of the four conserved residues (D34, E35, D141 and D169) which form the active site of RNASEH2A are indicated with red lines (based on murine Rnaseh2a (Shaban *et al.*, 2010)). The B subunit of RNase H2 has 2 isoforms: the full length protein (RNASEH2B) containing a PCNA-interacting protein (PIP) box motif at amino acid residues 294 – 301 (Chon *et al.*, 2009), and a shorter alternative isoform (RNASEH2Balt) which lacks the PIP box (Martin Reijns, MRC HGU, personal communication).

Both *Saccharomyces cerevisiae* (*S. cerevisiae*) and human RNASEH2B/C can form complexes independent of the catalytic A subunit (Chon *et al.*, 2009; Rohman *et al.*,

2008). The smaller C subunit interacts with the N-terminal region of the B subunit (as indicated in Figure 1.3). This heterodimer structurally resembles the subdomains of the eukaryotic transcription factor IIF (TFIIF) complex (Shaban *et al.*, 2010). The B/C subcomplex may bind proteins other than RNASEH2A, probably via the C-terminus of RNASEH2B. This could support a function for these proteins supplementary to their role as components of the enzymatically active RNase H2 complex (Chon *et al.*, 2009; Jeong *et al.*, 2004; Shaban *et al.*, 2010).

RNASEH2B was identified as being acetylated in a high-throughput screen (Choudhary *et al.*, 2009); however post-translational modifications of the A and C subunits have not been reported.

1.1.4.2. *In vivo* functions of the RNase H enzymes

The functions of the RNase H enzymes *in vivo* have not yet been fully characterised, but there is strong evidence to suggest that mitochondrial RNase H1 plays an essential role in mammalian embryogenesis (Cerritelli *et al.*, 2003) (discussed in Section 1.1.5). The nuclear isoform of RNase H1 and the predominantly nuclear RNase H2 complex may function during transcription (Broccoli *et al.*, 2004) (discussed in Section 1.3.2).

The ability of RNase H2 to hydrolyse 5′ of single ribonuleotides embedded in dsDNA indicates that it may have a role in the recognition and removal of these misincorporated bases from genomic DNA (Arudchandran *et al.*, 2000; Eder and Walder, 1991; Eder *et al.*, 1993). The misincorporation of ribonuleotides into DNA can potentially occur during DNA synthesis (Joyce, 1997), especially when an error-prone DNA polymerase is utilised (Bergoglio *et al.*, 2003; Nick McElhinny *et al.*, 2010). The presence of a ribose in DNA prevents it from adopting the normal dsDNA conformation and instead results in a distortion of the DNA backbone (Eder *et al.*, 1993). The progression of DNA polymerase can be blocked as the active site of the enzyme is unable to accommodate the extra oxygen molecule in the 2′ position in the ribose ring, stalling DNA replication (Nick McElhinny *et al.*, 2010). RNase H2

may be required to excise these bases during DNA replication and repair, which potentially provides an explanation for the highly conserved nature of the RNase H2 enzymes and their representation throughout all kingdoms of life.

1.1.4.3. Interaction partners of the RNase H2 subunits

Yeast two-hybrid experiments have identified only one protein other than the B and C subunits which interacts with the A subunit *S. cerevisiae* RNase H2, the Rap1p-interacting factor, Rif2p (Jeong *et al.*, 2004). Following its recruitment to telomeres by Rap1p, Rif2p forms a complex which regulates telomere length (Wotton and Shore, 1997). The B and C subunits of yeast RNase H2 are not reported to interact with Rif2p.

The RNASEH2B protein contains a proliferating cell nuclear antigen (PCNA) interacting protein box (PIP-box) at the C-terminus (Figure 1.3), highly conserved among different species (Chon *et al.*, 2009; Crow *et al.*, 2006b). This ability to interact with PCNA indicates a role for RNase H2 in DNA replication and repair (Chon *et al.*, 2009). In mouse and human, a protein-coding splice variant of *RNASEH2B* exists. The isoform, RNASEH2Balt, lacks the PIP-box (Figure 1.3.) but still forms an enzymatically active complex with the A and C subunits *in vitro* (Martin Reijns, MRC HGU, personal communication). It is not yet clear if this alternative complex exists *in vivo* and if it has a specialised function distinct from that of the known RNase H2 complex.

Interaction studies using the B subunit of *S. cerevisiae* RNase H2 suggest that it is also capable of interacting with additional proteins (Jeong *et al.*, 2004), including the Cdk-activating kinase Cak1p, which is involved in cell cycle progression through S-phase and G2/M-phase (Sutton and Freiman, 1997) and the transcriptional regulator forkhead protein 1 (Fkh1) which plays a key role in yeast mating-type switching (Sun *et al.*, 2002).

Known protein binding domains have not yet been identified in the C subunit of the RNase H2 complex. However, high-throughput interaction studies in *S. cerevisiae* suggest that it could interact with members of the MAP kinase Kss1p complex, which is activated prior to mating to induce G1-arrest (Cherkasova *et al.*, 1999). Additionally, an interaction with Topoisomerase II (Top2p), which functions in replication and repair pathways (Holm *et al.*, 1985), and has recently been implicated in the suppression of R loop formation (El Hage *et al.*, 2010) has also been suggested (Ho *et al.*, 2002).

1.1.5. RNase H-deficient eukaryotic cells

1.1.5.1. Yeast deficient in RNase H activity are viable

S. cerevisiae cells with a deletion of the entire *RNH1* gene are viable and show only a slight increase in sensitivity to the DNA-damaging agent ethyl methansuplhonate (EMS) (Arudchandran *et al.*, 2000). Total RNase H activity is not reduced in these cells, attributable to an increased level of RNase H2 activity which indicates that RNase H2 can compensate for a loss of RNase H1 activity in yeast (Arudchandran *et al.*, 2000). Cells lacking RNase H1 activity also exhibit normal mtDNA replication (Arudchandran *et al.*, 2000).

Deletion of the gene encoding the A subunit of RNase H2 (*RNH2A*) does not affect viability but does result in a 75% decrease in the total RNase H activity in these cells compared to wild type (Arudchandran *et al.*, 2000; Frank *et al.*, 1998b). RNase H activity was also detected in cells with a double deletion of both *RNH1* and *RNH2A*, which although viable, have an increased sensitivity to treatment with either hydroxyurea (HU) or EMS (Arudchandran *et al.*, 2000). HU inhibits the enzyme ribonucleoside diphosphate reductase, depleting the dNTP pool available during S-phase of the cell cycle, and causing cells to undergo slow DNA replication. Cells lacking RNase H1 respond as wild-type to HU treatment, but the co-deletion of *RNH2A* prolongs the S-phase arrest and eventually results in cell death (Arudchandran *et al.*, 2000). This suggests RNase H2 is required during DNA

replication in S-phase, and possibly in the repair of HU-mediated DNA damage during the G2/M-phase of the cell cycle (Arudchandran *et al.*, 2000).

1.1.5.2. *Rnaseh1*-null mice are embryonic lethal

In contrast to prokaryotes and yeast which are viable when deficient in RNase H1 activity, RNase H1-null mice die *in utero* (Cerritelli *et al.*, 2003). Initially, the embryos develop normally but they have a greatly reduced mtDNA content; approximately 10% of wild-type at 7.5 days post coitum (d.p.c.) (E7.5). This leads to developmental arrest at E8.5 and suggests an essential role for RNase H1 in the replication of mtDNA during development (Cerritelli *et al.*, 2003). Mouse models defective in RNase H2 activity have yet to be fully characterised and published.

1.1.5.3. Mutations in the genes encoding human RNase H2 cause Aicardi-Goutières syndrome (AGS)

Mutations in the human *RNASEH1* gene have not been reported, however mutations in any of the three genes encoding the human RNase H2 complex subunits can cause the childhood inflammatory disorder Aicardi-Goutières syndrome (AGS) (Crow *et al.*, 2006b). This is the first example of an RNase H enzyme being implicated in a human disease. Understanding the role of the enzyme in disease pathogenesis should also provide an opportunity to fully elucidate the *in vivo* functions of RNase H2.

1.2. Aicardi-Goutières syndrome (AGS)

AGS is a rare neuroinflammatory disease, first described in 1984 by Jean Aicardi and François Goutières (Aicardi and Goutieres, 1984). They reported eight children from five families with early-onset progressive famililal encephalopathy, basal ganglia calcification and chronic cerebrospinal fluid (CSF) lymphocytosis. These features closely mimicked an intrauterine viral infection, but TORCH (toxoplasmosis, rubella, cytomegalovirus and herpes simplex virus types 1 and 2) investigations proved negative. This, together with the recurrence of the disease in families, suggested that there could be a genetic basis for disease (Aicardi and Goutieres, 1984). Subsequent confirmation of this has led to the description of AGS as a Mendelian mimic of congenital viral infection (Crow *et al.*, 2006b; Crow and Livingston, 2008).

The clinical phenotype of AGS, in particular the notable overlap with the common autoimmune disease systemic lupus erythematosus (SLE), is described in more detail in this section.

1.2.1. The clinical phenotype of AGS

At disease onset, AGS patients develop an acute onset encephalopathy and often display irritability, inconsolable crying, feeding difficulties, dystonic posturing, recurrent fevers and a loss of social and motor skills (Crow and Livingston, 2008; Lanzi *et al.*, 2002; Rice *et al.*, 2007b). These periods usually last several months, before the condition stabilises and there is no further disease progression (Rice *et al.*, 2007b). The majority of patients are left severely physically and mentally impaired, with death often occurring during childhood (Goutieres, 2005; Rice *et al.*, 2007b).

The phenotype of AGS has been shown to be variable in severity, even between affected siblings (Ostergaard *et al.*, 1999; Ramantani *et al.*, 2010; Rice *et al.*, 2007b; Verrips *et al.*, 1997) and so a range of diagnostic criteria has been established (reviewed by Orcesi *et al.*, 2009). The majority of AGS patients do not present with the disease at birth and are born following a normal pregnancy (Goutieres *et al.*,

1998; Rice *et al.*, 2007b). However, a subgroup of patients are affected at birth, exhibiting an abnormal neurology, jitteriness and seizures in addition to extraneurological features including enlarged spleens and reduced levels of platelets (Crow and Livingston, 2008; Menson and Lyall, 2005; Rice *et al.*, 2007b). This presentation considerably overlaps with that of congenital infection (Table 1.1).

The other patients present at variable times beyond the first few days of life, some developing disease progressively following several months of normal development (Aicardi and Goutieres, 1984; Crow and Livingston, 2008; Rice *et al.*, 2007b). The trigger for disease onset is unknown (Crow and Livingston, 2008) but extensive testing for the presence of a range of viruses in AGS patients using a variety of techniques has invariably proved negative (Goutieres, 2005; Lebon *et al.*, 2002).

Table 1.1. A comparison of the overlapping clinical features of AGS, the common adult-onset autoimmune disease systemic lupus erythematosus (SLE) and congenital viral infection.

		AGS	SLE	Congenital viral infection
Origin		Genetic (monogenic)	Multifactorial: complex genetic and environmental factors	Viral infection
Age at presentation		From birth to approx. 12 months	Generally late teens - 40 years of age	At birth
Neurological features	Basal ganglia calcification	+	+	+
	Destruction of white matter	+	+	+
	Microcephaly	+	-	+
Immunological features	Elevated levels of CSF IFN-α	+	+	+
	Hypergammaglobinemia	+	+	+
	Autoantibodies	+/-[a]	+	-
Systemic features	Immunoglobulin deposits in skin[b]	+	+	-
	Thrombocytopenia	+	+	+
	Haemolytic anaemia	+[c]	+	+
	Hepatosplenomegaly	+	-[d]	+
	Fever	+	+	+

SLE is a multisystem disorder with a wide range of clinical features, only a small subset of which overlap with AGS. Similarly, congenital viral infection has pleiotropic manifestations that are absent in AGS. [a]The presence of autoantibodies has been reported in a small proportion of AGS patients. [b]Chilblain-like lesions with IgM deposition. [c]Reported in a small number of AGS patients. [d]Although splenomegaly can be present. Table reproduced from Rigby *et al*. 2008.

Neuroimaging of AGS patients showed they have microcephaly, regions of intracellular calcification and white matter destruction (Figure 1.4) (Bale *et al.*, 1985; Crow *et al.*, 2006b; Numazaki and Fujikawa, 2003; Rice *et al.*, 2007b). The calcifications, which are typically in the basal ganglia, can vary greatly in size and number, although this does not correlate with the severity of disease status (Goutieres, 2005; Ostergaard *et al.*, 1999) and they tend to remain stable over time from the initial diagnosis of AGS (Lanzi *et al.*, 2002; Lanzi *et al.*, 2005). A similar neuropathology is seen in congenital viral infection (Bale *et al.*, 1985; Belman *et al.*, 1986; Belman *et al.*, 1996) and can occur in SLE patients (Figure 1.4) (Raymond *et al.*, 1996).

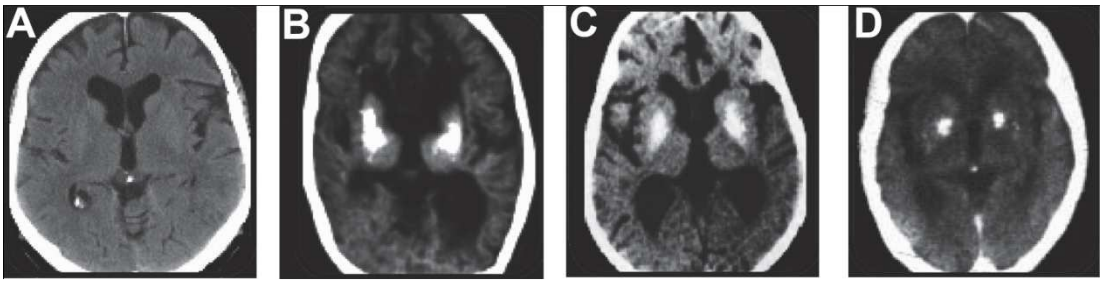


Figure 1.4. The neuropathology of AGS patients is reminiscent of both congenital viral infection and autoimmune disease.
Axial CT scans of normal (A), AGS (B), congenital HIV (C) and SLE (D) brain showing calcification of basal ganglia (central regions of high signal, B-D) and white matter destruction. (Rigby *et al.*, 2008)).

1.2.2. Immunological phenotype of AGS

In addition to lymphocytic infiltration of the CSF, AGS patients display a number of other immunological phenotypes (Rice *et al.*, 2007b), some of which are listed in Table 1.1. However, the immunological profile of AGS also shows variability between patients.

The similarity of the AGS pathogenesis to that of intrauterine viral infections prompted measurement of the antiviral cytokine interferon alpha (IFN-α) in patients (Lebon *et al.*, 1988). Generally, AGS patients have levels of CSF IFN-α greatly exceeding the normal range of less than 2 International Units (IU)/ml (Goutieres *et al.*, 1998), however, the levels tend to fall as the patients age as does the white cell count of the CSF (Lebon *et al.*, 2002; Rice *et al.*, 2007b). The average levels of IFN-α are higher in patients with early-onset AGS compared with patients who present with the disease after 3 months of age (Lebon *et al.*, 2002). The biological significance of elevated levels of IFN-α is discussed further in Section 1.5 of this thesis.

Elevated levels of CSF pterins, a biomarker for CNS inflammation (Dale *et al.*, 2009), are also found in AGS patients (Blau *et al.*, 2003; Rice *et al.*, 2007b; Wassmer *et al.*, 2009). Extraneurological inflammatory features include chilblain-like lesions at the extremities resulting from an inflammatory vasculopathy, hepatosplenomegaly and thrombocytopenia (reviewed by Aicardi, 2002; Crow and Livingston, 2008; Goutieres, 2005).

1.2.3. Cytokine production in AGS

Elevated levels of IFN-α can be detected in the blood of some patients, but levels are usually lower than the equivalent found in the CSF, indicative of predominantly intrathecal production (Goutieres, 2005; Lebon *et al.*, 1988; Lebon *et al.*, 2002; van Heteren *et al.*, 2008). Consistent with this, glial cells, specifically astrocytes, have been identified as producing cytokines, most notably IFN-α, in AGS patients (van Heteren *et al.*, 2008). Astrocytes are also the source of cytokine production in viral-induced encephalopathies (van Heteren *et al.*, 2008). The CNS can direct its own immune response by transforming resident cells into components of the immune response and the role of astrocytes (the most prevalent cell in the CNS) in the induction of an inflammatory response has been well characterised (Carpentier *et al.*, 2005).

A study of 30 cytokines and growth factors in the CSF and sera of a cohort of AGS patients showed that in addition to IFN-α, levels of the chemokines CXCL10 (IP-10) and CCL2 (MCP-1) are significantly elevated in the CSF of AGS patients (van Heteren *et al.*, 2008). Similarly, raised levels of these chemokines were detected in the CSF of patients with a known viral infection of the brain (van Heteren *et al.*, 2008). In all cases, astrocytes were responsible for their production. Increased amounts of CXCL10 and CCL2 were also detected in the sera of AGS patients, although levels were lower than in the CSF (van Heteren *et al.*, 2008).

CXCL10 is likely to be produced as a secondary effect of IFN-α (Campbell *et al.*, 1999; van Heteren *et al.*, 2008) as it is known to be induced by type I and type II IFNs *in vitro* (Luster *et al.*, 1985; Padovan *et al.*, 2002). The receptor for CXCL10, CXCR3 (Loetscher *et al.*, 1998), is expressed on a number of immune cells, particularly on activated Th1 cells (reviewed by Lacotte *et al.*, 2009), therefore it is possible that CXCL10 in the CSF of AGS patients is responsible for the accumulation of lymphocytes there (Klein, 2004; van Heteren *et al.*, 2008). This is a central element of disease pathogenesis in inflammatory disorders of the CNS, including viral encephalopathies (Rosler *et al.*, 1998) and the autoimmune disease multiple sclerosis (Sorensen *et al.*, 2002). A study of astrocyte-targeted overproduction of CXCL10 in mice showed that whilst the mice had leukocyte infiltrates into the brain, mainly comprising of neutrophils and T cells, they did not exhibit any pathological signs of an inflammatory disease of the CNS (Boztug *et al.*, 2002). This suggests that upregulated expression of CXCL10 in the CNS can direct lymphocytes to the blood-brain barrier (BBB) but additional signals are required to allow the crossing of the BBB and subsequent induction of neuronal damage and disease pathology (Boztug *et al.*, 2002; Klein, 2004; Trifilo and Lane, 2003). Astrocyte-targeted chronic expression of CCL2 resulted in neurological impairment due to perivascular lymphocytosis and impairment of the BBB (Huang *et al.*, 2005). Therefore, it is possible that production of high levels of both of these chemokines is responsible for the CSF lymphocytosis in AGS.

Levels of the interleukins IL-6, IL-8 and of vascular endothelial growth factor (VEGF) are significantly increased in the CSF of patients with viral infections but not in the CSF of AGS patients (van Heteren *et al.*, 2008). This clearly distinguishes the AGS cytokine profile from that of viral infection of the brain, suggesting a different mechanism of disease onset and pathology.

1.2.4. Phenotypic overlap with systemic lupus erythematosus (SLE)

As mentioned previously in this section, a number of the clinical and immunological features displayed in AGS patients can also occur in the common adult autoimmune disease SLE (Table 1.1), notably the elevated level of IFN-α in the CSF. The co-existence of AGS and SLE has been identified in two patients (one with a mutation in the *TREX1* gene and the other in *SAMHD1*; these genes are discussed in more detail in Section 1.2.5) (Ramantani *et al.*, 2010), suggesting a common mechanism of disease pathogenesis is shared by the two disorders.

SLE is a chronic, remitting and relapsing disorder disproportionally affecting females. It has a prevalence of approximately 0.1% of the general population, although this can vary with different ethnic populations (D'Cruz *et al.*, 2007). SLE affects multiple systems of the body and patients exhibit a wide range of clinical presentations which can include arthritis, fever, skin rashes, photosensitivity, neuroinflammation and kidney dysfunction often leading to renal failure. Intracranial calcifications can be seen in small proportion of SLE patients (Figure 1.4) (Raymond *et al.*, 1996).

The aetiology of SLE is not yet fully understood; genetic and environmental factors are known to be involved in the development of disease (D'Cruz *et al.*, 2007; Vyse and Kotzin, 1998). Increased susceptibility to SLE has been associated with polymorphisms in a wide range of genes including; *TREX1* (discussed further in Section 1.2.5); DNase I (*DNASE1*) (discussed in 1.4.1); interferon regulatory factor 5 (*IRF5*) (Graham *et al.*, 2006; Kozyrev *et al.*, 2007); protein tyrosine phosphatise N22

(*PTPN22I*) (Kyogoku *et al.*, 2004); programmed cell death 1 gene (*PDCD1*) (Prokunina *et al.*, 2002); and deficiencies in the early complement components, C1q, C1r, C1s, C4 and C2 (Manderson *et al.*, 2004).

1.2.5. The genetics of AGS

A genome-wide linkage analysis study involving 13 AGS families established linkage of the disease to chromosome 3p21 (the *AGS1* locus) in approximately 50% of these families (Crow *et al.*, 2000). Further genome-wide analysis of 10 families incompatible with linkage to chromosome 3p21 identified a second AGS locus (*AGS2*) at chromosome 13q14-21 (Ali *et al.*, 2006), demonstrating that AGS was a single gene disorder with locus heterogeneity.

1.2.5.1. Mutations in the *RNASEH2* genes causes AGS

Genetic refinement of the *AGS2* locus identified *FLJ11712/DLEU8* as mutated in 18 families in a larger cohort of patients (Crow *et al.*, 2006b). Bioinformatics analysis revealed homology with a distant *S. cerevisiae* orthologue, *RNH202*, which encodes an essential component of the yeast RNase H2 complex, Rnh2Bp (Crow *et al.*, 2006b; Jeong *et al.*, 2004).

A third AGS locus (*AGS3*) at chromosome 11q13.2 was identified and refined following the discovery of an ancestral haplotype in five families, which contained the *AYP1* gene (Crow *et al.*, 2006b). AYP1 was known to co-purify with RNASEH2A, the human orthologue of yeast Rnh2Ap (Frank *et al.*, 1998c; Jeong *et al.*, 2004). Sequence analysis subsequently identified *AYP1* as the AGS-causative gene at the *AGS3* locus (Crow *et al.*, 2006b). It was confirmed to encode the human orthologue of *S. cerevisiae* Rnh2Cp and was consequently renamed *RNASEH2C* (Crow *et al.*, 2006b).

Sequencing of *RNASEH2A* on chromosome 19p13.13 in an AGS family who did not show linkage to the *AGS1-3* loci identified mutations in two patients (Crow *et al.*, 2006b). Subsequent experiments confirmed that the three proteins, RNASEH2A,

RNASEH2B and RNASEH2C, interact *in* vitro and form an enzymatically active complex with Type 2 RNase H activity (Crow *et al.*, 2006b).

1.2.5.1.1. *RNASEH2* mutations

The majority of AGS causing mutations identified in the *RNASEH2* genes are hypomorphic missense mutations inherited in an autosomal recessive pattern (Crow *et al.*, 2006b; Rice *et al.*, 2007b). Parents are obligate carriers for the mutations (Crow *et al.*, 2006b; Rice *et al.*, 2007b). Biallelic null mutations have not been reported in any of the *RNASEH2* genes, suggesting that the resulting phenotype would be lethal *in utero* (Rice *et al.*, 2007b).

The most comprehensive study of AGS-causing mutations to date, involving a cohort of 127 families, identified biallelic mutations in all RNase H2-AGS patients with the exception of three families with a single *RNASEH2B* mutation and one family with a single *RNASEH2A* mutation (Rice *et al.*, 2007b). Mutations in *RNASEH2A* are rare and consist predominantly of missense mutations, located throughout the gene. The *RNASEH2B* A177T change is a recurrent mutation, seen in almost every AGS patient with a *RNASEH2B* mutation. No mutations were identified in the final three exons of the gene (exons 9, 10 and 11), which is the region that can undergo alternative splicing to produce RNASEH2Balt. All *RNASEH2C* mutations identified are homozygous missense mutations, found throughout the gene with a recurring R69W mutation present in a large proportion of patients (Crow *et al.*, 2006b; Rice *et al.*, 2007b).

1.2.5.1.2. The effect of *RNASEH2* mutations on protein function

Recombinant RNase H2 complexes with AGS-causing mutations in the B and C subunits displayed normal enzymatic activity *in vitro* (Perrino *et al.*, 2009). Therefore, it was hypothesised that altered RNase H2 activity is not responsible for the pathogenesis of AGS, but instead arises as a consequence of alternations in the assembly, stability and localisation of the complex (Chon *et al.*, 2009; Perrino *et al.*, 2009). However, this is not true *in vivo*, as the RNase H2 activity of lymphoblastoid

cell lines (LCLs) derived from AGS patients is significantly reduced (Martin Reijns, MRC HGU, personal communication), indicating that AGS is a consequence of reduced RNase H2 enzymatic activity.

1.2.5.2. Mutations in *TREX1* cause AGS

The identification of AGS-causing mutations in genes encoding an RNase prompted the sequencing of the single exon *TREX1* gene; the only gene in the *AGS1* locus known to encode a nuclease. *TREX1* mutations were subsequently identified in 10 AGS families (Crow *et al.*, 2006a).

1.2.5.2.1. The nuclease activity of TREX1

TREX1 encodes the 33kDa Three Prime Repair Exonuclease 1 (TREX1) protein, formerly known as DNase III, which constitutes the major 3′ → 5′ exonuclease activity in mammalian cells (Lindahl *et al.*, 1969; Mazur and Perrino, 1999). TREX1 and its close homologue TREX2, with which TREX1 shares 44% amino acid sequence identity (Mazur and Perrino, 1999), have no orthologues in lower eukaryotes, including yeast (Hoss *et al.*, 1999). However TREX1 does shares sequence homology with the *E.coli* proofreading protein DnaQ, a component of the DnaQ/MutD DNA polymerase III holoenzyme. Consequently, TREX1 was originally hypothesised to have a similar role in editing mismatched bases inserted during gap filling in DNA base excision repair by the error-prone DNA polymerase Pol β (Hoss *et al.*, 1999). However, as discussed in this section, this does not appear to be the case *in vivo* (Morita *et al.*, 2004).

TREX1 is a 314 amino acid protein comprising three distinct conserved sequence motifs (ExoI, ExoII and ExoIII) (Bernad *et al.*, 1989; Blanco *et al.*, 1992) (Figure 1.5). These Exo domains bind the divalent cation cofactors (Mg^{2+} or Mn^{2+}) required for enzymatic activity (Lindahl *et al.*, 1969; Mazur and Perrino, 1999). Unique amongst related exonucleases, TREX1 contains an 8 amino acid polyproline II helix which may provide a flexible region for interactions with other proteins (Brucet *et al.*, 2007; de Silva *et al.*, 2007), although to date no interaction partners for TREX1

have been identified. The C terminus of TREX1 is absent from TREX2 and other members of the DnaQ family (Hoss *et al.*, 1999; Mazur and Perrino, 1999; Mazur and Perrino, 2001a). It contains a leucine-rich transmembrane domain (TMD) which is responsible for tethering the protein to the endoplasmic reticulum (ER) (Lee-Kirsch *et al.*, 2007b; Richards *et al.*, 2007), presumed to be essential for the biological function of TREX1 (Lindahl *et al.*, 2009).

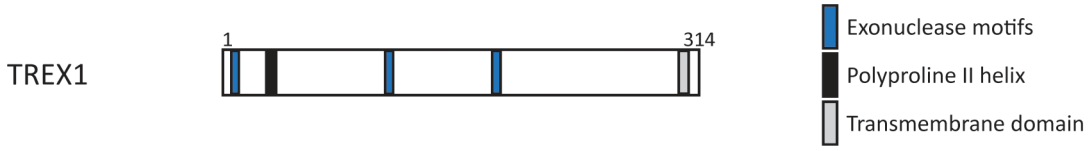


Figure 1.5. The human TREX1 protein.
Schematic diagram representing human TREX1. TREX1 contains three conserved Exonuclease domains (I, II and III) located between residues 8 – 14, 111 – 120 and 185 – 194. The N terminus of the protein also contains a polyproline II domain between amino acid residues 54 – 62, whilst the C terminus contains a transmembrane domain (residues 293 – 307).

Active as a homodimer, TREX1 catalyses the excision of nucleoside monophosphates from the 3′ terminus of DNA (Figure 1.6), having 3-4 fold greater activity on single-stranded DNA (ssDNA) than double-stranded DNA (dsDNA) (Lindahl *et al.*, 1969; Mazur and Perrino, 1999; Mazur and Perrino, 2001a). The enzyme has no activity on single-stranded RNA (ssRNA) or partial RNA:DNA hybrids and completely lacks endonuclease activity (Lindahl *et al.*, 1969; Mazur and Perrino, 1999).

The *TREX1* transcript is ubiquitously expressed (Mazur and Perrino, 2001b). The majority of the protein is located in the cytoplasm, but relocates to the nucleus in response to granzyme A activity (Chowdhury *et al.*, 2006) and genotoxic stress (Christmann *et al.*, 2010; Yang *et al.*, 2007). Nuclear TREX1 has a higher catalytic

activity than ER-associated TREX1, further suggesting that this localisation is connected to the biological function of the enzyme (Christmann *et al.*, 2010; Lindahl *et al.*, 2009).

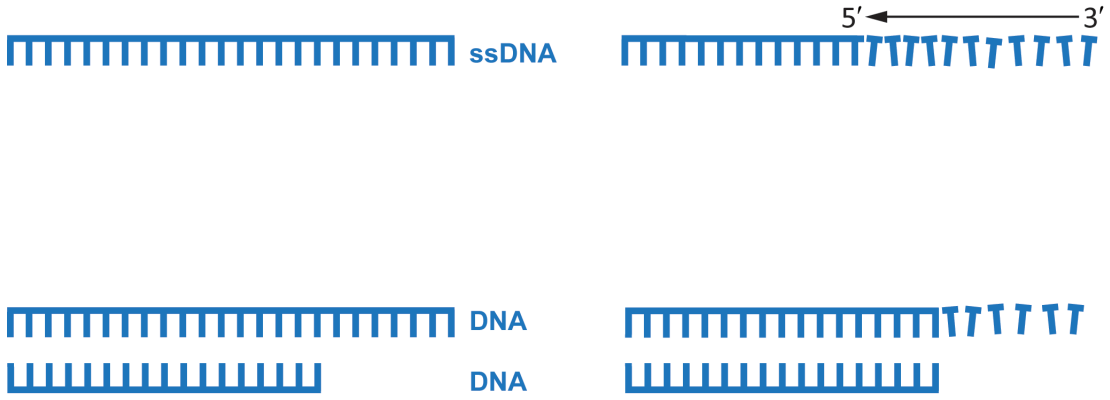


Figure 1.6. Preferred substrates of the 3' → 5' DNA exonuclease TREX1.
TREX1 excises single nucleotides from the 3' termini of ssDNA substrates, including 3' ssDNA overhangs, but can also cleave dsDNA (Lindahl *et al.*, 1969; Mazur and Perrino, 1999; Mazur and Perrino, 2001a)

TREX1 has been shown to be involved in a range of biological processes, predominantly by investigations using transgenic mice deficient in Trex1.

1.2.5.2.2. Trex1-deficient mice have an autoinflammatory phenotype

Trex1-homozygous null mice have significantly reduced 3′ → 5′ DNA exonuclease activity but are viable (Morita *et al.*, 2004). Similar to AGS patients, they do not have an increased incidence of cancer or spontaneous mutation frequency. However, they do develop a severe inflammatory myocarditis which leads to circulatory failure and a severely reduced lifespan (Morita *et al.*, 2004). Embryo transfer experiments confirmed that the lethal heart failure was not due to an increased susceptibility to cardiotoxic viruses (Morita *et al.*, 2004).

Post-mortem examination of *Trex1*$^{-/-}$ hearts showed that the heart muscle morphology was severely disturbed as a result of the lymphoid inflammatory infiltrate (Morita *et al.*, 2004). *Trex1*$^{-/-}$ hearts also had significantly increased expression levels of the type I interferon gene, *Ifnb1* (Stetson *et al.*, 2008). This provides further evidence to suggest that the *Trex1*$^{-/-}$ phenotype is due to inflammation in the absence of pathogenic agents, reminiscent of AGS. However, the mice lack the autoinflammatory neurological phenotype seen in AGS patients with loss-of-function mutations in *TREX1*.

Further investigations into the mechanism of autoimmunity in the *Trex1*$^{-/-}$ mice revealed that development of the inflammatory phenotype was mediated by innate immunity (Stetson *et al.*, 2008). Notably, mice deficient in Trex1 and the DNA recombinase RAG2 did not develop the lethal inflammatory cardiomyopathy but still had elevated levels of *IFN-β* expression, suggesting that type I IFN induction is not sufficient to cause autoimmune disease in *Trex*$^{-/-}$ mice without the involvement of the adaptive immune system (Stetson *et al.*, 2008).

1.2.5.2.3. Single-stranded DNA accumulates in TREX1-deficient cells

Primary fibroblast cells from AGS patients homozygous for *TREX1* loss-of-function mutations have an accumulation of DNA in their ER, which is eliminated by digestion with S1 nuclease (Yang *et al.*, 2007). A spontaneously transformed murine embryonic fibroblast (MEF) Trex1-null cell line also exhibited this phenotype (Yang *et al.*, 2007). Analysis of nucleic acid purified from the cytoplasm of these MEFs indicated that it was an ssDNA species approximately 60 – 65 nucleotides (nt) long (Yang *et al.*, 2007). Accumulating ssDNA was also found in the hearts of Trex1-deficient mice (Stetson *et al.*, 2008). Sequence analysis of this ssDNA revealed that much of it is derived from repetitive elements, and therefore Stetson *et al*. suggested that it is derived from endogenous retroviral DNA (Stetson *et al.*, 2008). Consequently, reverse transcribed DNA is a putative endogenous substrate for Trex1 (Stetson *et al.*, 2008).

1.2.5.2.4. Trex1 is a negative regulator of the innate immune response to DNA

Trex1 expression is induced *in vitro* in response to transfection with phosphodiester-containing dsDNA which lacks contiguous CpG motifs (Stetson *et al.*, 2008). This interferon stimulatory DNA (ISD) is known to induce type I IFN production in a Toll-like receptor (TLR)-independent mechanism, involving IRF3 activation (Stetson and Medzhitov, 2006). Trex1-deficient cells respond normally to ISD, indicating that Trex1 is not a cytosolic DNA sensor but instead is likely to function in the metabolism of endogenous ligands of the ISD pathway (Stetson *et al.*, 2008). Failure to do this results in the accumulation of unprocessed ssDNA and the development of lethal autoimmunity in mice deficient in Trex1 activity (Morita *et al.*, 2004; Stetson *et al.*, 2008).

1.2.5.2.5. TREX1 is a member of the SET complex

TREX1 is a component of the SET complex, an ER-associated multi-protein complex also containing the endonuclease NM23-H1, its inhibitor (the nucleosome assembly protein SET), together with HMBG2 and APE1 (reviewed by Chowdhury and Lieberman, 2008). The SET complex relocates to the nucleus in response to superoxide generated by mitochondria damaged by granzyme A protease (Chowdhury *et al.*, 2006). Granzyme A cleaves SET, which in turn releases NM23-H1 to introduce single-stranded nicks in the target cell. TREX1 then removes the resulting free 3′ ends, causing irreparable damage to the DNA, resulting in caspase-independent cell death (Chowdhury *et al.*, 2006). A *TREX1* mutation which affects enzyme activity and impairs this process has been identified in families with familial chilblain lupus, implicating TREX1 in disease pathogenesis (Lee-Kirsch *et al.*, 2007a; Rice *et al.*, 2007a).

The SET complex has also been implicated in the facilitation of HIV-1 infection by preventing the viral genome from autointegration: a suicidal mechanism whereby the viral genome accidentally inserts into itself (Yan *et al.*, 2009). During the generation of dsDNA from the ssRNA retroviral genome in the cytosol of infected cells,

unwanted byproducts such as autointegrated DNA and ssDNA are generated, which are bound to and degraded by the SET complex (Douville and Hiscott, 2010; Yan *et al.*, 2009). Failure to do this results in the accumulation of these defective nucleic acids which are unable to integrate into the host cell. It is hypothesised that subsequent detection of these DNA species by cytosolic receptors of the innate immune system induces a type I IFN antiviral response (Douville and Hiscott, 2010). Therefore, in the absence of the SET complex, autointegration is enhanced and HIV chromosomal integration is decreased (Yan *et al.*, 2009).

Similarly, it has also been demonstrated that TREX1 plays a direct role in the suppression of interferon production in response to HIV-1 infection by degrading viral reverse transcripts that would otherwise accumulate in the cytosol and trigger an innate immune response via an unknown cytosolic DNA sensor (Yan *et al.*, 2010). Together, these studies suggest that cellular TREX1 acts to promote HIV-1 replication although it is unclear why a host nuclease would have this function (Geijtenbeek, 2010).

1.2.5.2.6. TREX1 is inducible by certain types of DNA damage

Both human and murine TREX1 protein expression is induced in a time- and dose-dependent manner in response to treatment with UV and benzo(a)pyrene, which both introduce bulky lesions into DNA. This is completely dependent on the transcription factor AP-1 and coincides with the translocation of TREX1 to discrete nuclear foci (Christmann *et al.*, 2010), which also occurs in response to ionising radiation (Yang *et al.*, 2007). Additionally, *Trex1*-deficient cells are hypersensitive to treatment with UV and benzo(a)pyrene and show a delayed recovery from the resultant replication inhibition (Christmann *et al.*, 2010). They also show an abnormal response following exposure to ionising radiation (Yang *et al.*, 2007). These findings suggest that TREX1 may have a role in protecting the cell from certain DNA-damaging agents.

1.2.5.2.7. TREX1 and Aicardi-Goutières syndrome (AGS)

The majority of TREX1-AGS patients have biallelic null mutations that abrogate the enzymatic function of the protein (Crow *et al.*, 2006a; Rice *et al.*, 2007b), contrasting

with the predominantly missense mutations identified in the *RNASEH2* genes. *TREX1* mutations are found throughout the gene and the majority are inherited in an autosomal recessive manner, with all parents heterozygous for the allele (Rice *et al.*, 2007b). There are two reported cases of AGS resulting from different *de novo* heterozygous missense mutations, which alter the same amino acid residue in the ExoIII domain of TREX1 (Ramantani *et al.*, 2010; Rice *et al.*, 2007a). The $3' \rightarrow 5'$ DNA exonuclease activity on ssDNA in cells from one of the patients was comparable with controls (Rice *et al.*, 2007a). However, analysis of recombinant proteins has shown that this mutation reduces the ability of the enzyme to degrade nicked dsDNA (Lehtinen *et al.*, 2008). It also inhibits the activity of the wild type enzyme, thereby accounting for the dominant phenotype of the mutant allele (Lehtinen *et al.*, 2008). *TREX1* mutations constitute approximately 35% of AGS cases with a known mutation (Figure 1.8). The early-onset neonatal form of AGS is particularly prevalent amongst patients with *TREX1* mutations (Rice *et al.*, 2007b),

In addition to the *TREX1* mutations found in AGS patients, heterozygous mutations have been identified in patients with other inflammatory diseases, often sharing significant phenotypic overlap with AGS.

1.2.5.2.8. Familial chilblain lupus (FCL)

Autosomal dominant mutations in *TREX1* have been identified in patients with familial chilblain lupus, a subtype of lupus erythematosus which manifests in early childhood (Lee-Kirsch *et al.*, 2007a; Lee-Kirsch *et al.*, 2006; Rice *et al.*, 2007a). Patients develop ulcerating acral skin lesions, similar to the chilblain lesions seen in AGS, but do not develop a neurological phenotype (Lee-Kirsch *et al.*, 2006). One *TREX1* mutation found in patients with familial chilblain lupus has been shown to affect the dsDNA degradation activity of the enzyme, similar to the autosomal dominant *TREX1* mutation (Lehtinen *et al.*, 2008).

1.2.5.2.9. Systemic lupus erythematosus (SLE)

The study of a large cohort of SLE patients identified monoallelic *TREX1* mutations in approximately 3% of patients compared to only 0.1% of controls, suggesting that

there is a markedly increased relative risk of individuals with these *TREX1* variants developing SLE (Lee-Kirsch *et al.*, 2007b). Most of the mutations were missense mutations, located in the highly conserved C-terminal region (Figure 1.5). Additionally, two frameshift mutations, predicted to result in C-terminally truncated proteins, were demonstrated to have an altered subcellular localisation, although their enzymatic function was intact (Lee-Kirsch *et al.*, 2007b).

1.2.5.2.10. Retinal vasculopathy with cerebral leukodystrophy (RVCL)

Heterozygous frameshift mutations in *TREX1* which cause C-terminally truncated proteins were located in nine families with the autosomal dominant inflammatory disorder RCVL (Richards *et al.*, 2007). This adult-onset degenerative disease is characterised by visual loss, stroke and dementia, with death occurring 5 – 10 years from disease onset (Richards *et al.*, 2007). The mutant proteins retained full enzymatic activity but had an altered cellular distribution, predicted to affect their interaction with other proteins/substrates and consequently the normal biological function of the protein (Richards *et al.*, 2007).

The identification of *TREX1* mutations in human diseases other than AGS has shown that the nature and location of a mutation within the gene determine the consequences *in vivo*. Mutations completely abrogating the exonuclease activity of TREX1 result in severe, early-onset AGS, whilst heterozygous mutations which alter the subcellular localisation of the protein (leaving the enzymatic activity intact) result in the adult-onset disorders SLE and RCVL. This also illustrates the importance of TREX1 enzymatic activity at the correct location within the cell.

1.2.5.3. Mutations in *SAMHD1* cause AGS

A genotype-phenotype study of AGS patients in 2007 revealed that 17% of patients did not have mutations in *RNASEH2A, RNASEH2B, RNASEH2C* or *TREX1* (Rice *et al.*, 2007b). Genetic analysis of these patients located a fifth AGS locus (*AGS5*) on chromosome 20q11 (Rice *et al.*, 2009). Mutations in the *SAMHD1* gene, encoding the SAM domain and HD domain 1 protein (SAMHD1, also known as Dendritic

Cell-derived Interferon γ-induced Protein (DCIP)), were identified in 13 AGS families. However, the absence of mutations in the five AGS genes in a proportion of patients with clinically diagnosed AGS suggested that at least one other AGS locus exists (Ramantani *et al.*, 2010; Rice *et al.*, 2009).

1.2.5.3.1. *SAMHD1* mutations

SAMHD1 encodes a 72 kDa nuclear protein (Figure 1.7). As for the *TREX1* mutations reported in AGS patients, both biallelic null mutations and missense mutations were found in *SAMHD1*, affecting both the protein expression levels and nuclear distribution of the protein (Rice *et al.*, 2009).

1.2.5.3.2. The cellular function of SAMHD1

The *SAMHD1/DCIP* gene was originally identified as an orthologue of the murine *Mg11* gene (Li *et al.*, 2000). SAMHD1 is the only human protein to contain a sterile alpha motif (SAM) domain and a HD domain in tandem (Figure 1.7) (Rice *et al.*, 2009); however both domains are found separately in a range of proteins with diverse biological functions. SAM domains mediate interactions with other proteins, both SAM and non-SAM domain containing, and are capable of binding RNA hairpins and lipids (reviewed by Qiao and Bowie, 2005). HD domains are found in a superfamily of metal dependent phosphohydrolases, with predicted functions in nucleic acid metabolism and signal transduction (reviewed by Aravind and Koonin, 1998). However, SAMHD1 failed to interact with nucleic acid substrates including RNA stem loops, ssRNA, ssDNA and dsDNA, as assessed by surface plasmon resonance (Rice *et al.*, 2009), and so it is not yet clear if this protein can exhibit nuclease activity.

Figure 1.7. The human SAMHD1 protein.
Schematic diagram of SAMHD1. The sterile alpha motif (SAM) domain is situated between amino acid residues 42 – 110, and the HD domain extends from residues 164 – 319. Domain data sourced from the Pfam database (Finn *et al.*, 2010).

1.2.5.3.3. SAMHD1 and innate immunity

SAMHD1 is expressed widely, but is absent in normal human brain and thymus tissue (Liao *et al.*, 2008). The pro-inflammatory cytokine tumour necrosis factor (TNF)-α leads to the upregulation of SAMHD1 expression in human lung fibroblasts in an IRF1-dependent manner (Liao *et al.*, 2008). Similarly, murine *SAMHD1* expression is upregulated in a model of acute lung injury and may also contribute to Th1-mediated immune responses (Liao *et al.*, 2008).

In macrophages, *SAMHD1* expression increases in response to ISD in a type I IFN-dependent manner (Rice *et al.*, 2009). This response is TLR-independent, and suggests that similar to TREX1, SAMHD1 may function in preventing the activation of innate immunity by self-nucleic acids within the cell (Rice *et al.*, 2009). A similar role for RNase H2 has not yet been demonstrated, nor is it known if RNase H2 substrates accumulate in RNase H2-deficient cells and trigger autoimmunity.

1.2.5.4. Spectrum of mutations in AGS patients

The majority of identified mutations in AGS patients are in the *RNASEH2* genes, in particular *RNASEH2B* (Figure 1.8). Onset of disease in AGS patients with *RNASEH2B* mutations is significantly later than in patients with mutations in *RNASEH2A*, *RNASEH2C* or *TREX1* (Rice *et al.*, 2007b). RNASEH2B-AGS patients often exhibit a milder phenotype than other AGS patients and are significantly more likely to survive into early adulthood (Rice *et al.*, 2007b), suggesting that a possible genotype-phenotype correlation exists within AGS.

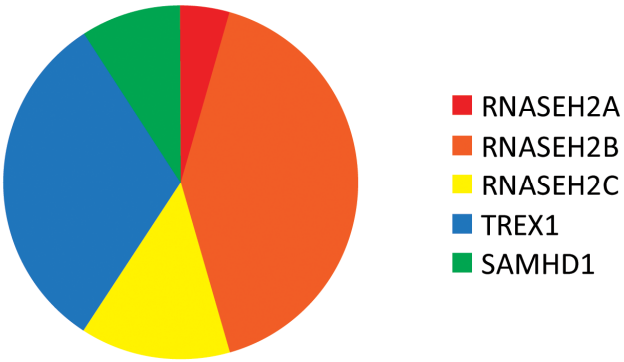


Figure 1.8. The proportion of AGS families with mutations in each of the five known AGS genes.
The majority of AGS families have mutations in the genes encoding the RNase H2 complex (45% have mutations in *RNASEH2B*, 15% in *RNASEH2C* and <5% in *RNASEH2A*). 35% of families have mutations in *TREX1*, whilst approximately 10% have mutations in the recently identified *SAMHD1* gene. (Compiled using data from Crow and Rehwinkel, 2009).

1.3. Nucleases in autoimmunity and infection

AGS is not the only human disease caused by defective nuclease activity. The DNases I and II also function to prevent accumulation of their nucleic acid substrates; failure to do so results in the development of multisystem autoimmune disease. However, nucleases can also play a direct role in the innate immune response, as exemplified by RNase L (as discussed in Section 1.3.3). Therefore, there exist a number of different ways by which the AGS nucleases might act to cause disease. This section describes in more detail the development of human disease as a consequence of mutations in genes encoding nucleases.

1.3.1. DNase I and SLE

DNase I is the major nuclease present in serum and urine (Nadano *et al.*, 1993). It acts on ssDNA, dsDNA and chromatin, preferentially cleaving at phosphodiester bonds adjacent to a pyrimide base, resulting in products with 5′-phosphate and 3′-hydroxyl (OH) ends. DNase I activity is reduced in the serum of SLE patients but not in patients with the other common autoimmune diseases rheumatoid arthiritis and scleroderma (Chitrabamrung *et al.*, 1981). Enzyme levels correlate with disease severity; patients with active lupus, in particular active renal involvement, exhibit the lowest levels of DNase I activity (Chitrabamrung *et al.*, 1981).

Heterozygous mutations in *DNASE1* correlating with reduced enzymatic activity have been described in a small number of individuals with SLE (Bodano *et al.*, 2006; Shin *et al.*, 2004; Yasutomo *et al.*, 2001). Additionally, autoantibodies against DNase I were identified in the sera of a significant number of SLE patients and were shown to inhibit enzymatic activity (Yeh *et al.*, 2003). Murine models of SLE confirmed the association with DNase I; SLE-prone NZB/NZW mice have decreased serum DNase I activity (Macanovic and Lachmann, 1997), and *Dnase1*$^{-/-}$ mice generated by gene targeting exhibit a phenotype very similar to SLE, including the development of anti-nuclear antibodies (ANA) (Napirei *et al.*, 2000).

The antibodies characteristic of SLE can mediate organ damage by binding to host tissues. Together with autoantigens, they can form immune complexes (ICs) which deposit in vascular systems, leading to inflammation, tissue destruction and the subsequent release of self-nucleic acids (reviewed by Chan *et al.*, 1999; Singh, 2005). Autoantibodies forming ICs with DNA derived from dying cells have been shown to activate innate immunity in plasmacytoid dendritic cells (pDCs) via engagement with Toll-like receptor 9 (TLR9) and FcγRIIa, mediated by High Mobility Group Box 1 (HMGB1) and Receptor for Advanced Glycation End-products (RAGE) (Barrat *et al.*, 2005; Lovgren *et al.*, 2004; Means *et al.*, 2005; Ronnblom and Alm, 2003; Tian *et al.*, 2007). Therefore, failure to clear debris from apoptotic or necrotic cells can result in the activation of autoimmune recognition by the innate immune system. DNase I, in cooperation with the complement factor C1q, plays a key role in the degradation of chromatin released from dying cells (Gaipl *et al.*, 2004). Consistent with this, deficiencies in chromatin clearance are seen both in human SLE patients and mouse models of the disease, which cannot be attributed to defects in populations of scavenger cells such as macrophages and immature DCs (Licht *et al.*, 2004; Munoz *et al.*, 2009).

Novel mutations in *DNASEI* have been identified in a small cohort of patients with autoimmune thyroid disease (AITD). These mutations are predicted to alter the stability of the enzyme and consequently, these patients have reduced serum DNase I activity (Dittmar *et al.*, 2008). This suggests that DNase I may be implicated in the pathogenesis of other autoimmune diseases.

1.3.2. DNase II and inflammatory disease

DNase II (formerly known as acid DNase) non-specifically cleaves both ssDNA and dsDNA, resulting in products with a 3′ phosphate (Baker *et al.*, 1998; Bernardi, 1968). It does not require Mg^{2+} for activity and displays optimal catalytic activity at an acidic pH (Baker *et al.*, 1998; Bernardi, 1968). DNase II localises to the lysosomes of macrophages (Kawane *et al.*, 2001).

A polymorphism in the human *DNASEII* gene resulting in decreased enzymatic activity has been identified (Yasuda *et al.*, 1992). However, having reduced DNase II activity in itself is not pathogenic, as people with the lower activity allele were healthy (Yasuda *et al.*, 1992).

DNase II has been reported to cleave the G-rich regions within the Ig switch recombination sequences (Evans and Aguilera, 2003; Lyon *et al.*, 1996), however studies in mice have shown that the primary function of DNase II is the degradation of DNA engulfed by macrophages (Kawane *et al.*, 2001).

The extensive analysis of *DNase II* null mice, which die late in embryonic development, revealed an essential role for DNase II in degrading the DNA of apoptotic cells and nuclei expelled from erythroid precursor cells phagocytosed by macrophages (Kawane *et al.*, 2001; Kawane *et al.*, 2003). Consequently, embryos lacking DNase II activity exhibit defects in erythropoesis in the liver, leading to severe anaemia (Kawane *et al.*, 2001; Krieser *et al.*, 2002), in addition to impaired thymic development (Kawane *et al.*, 2003). Fragments of self-DNA were shown to accumulate in the macrophages of these embryos (Kawane *et al.*, 2003; Krieser *et al.*, 2002). The presence of elevated levels of the cytokines IFN-β, IFN-γ and TNF-α in the thymii of *DNase II*$^{-/-}$ embryos implied that the innate immune system becomes activated in response to the undegraded DNA engulfed by the macrophages (Kawane *et al.*, 2003). IFN-β in particular is reported to be cytotoxic to erythroblasts and lymphocytes (Yoshida *et al.*, 2005). The IFN production in *DNase II*$^{-/-}$ mice was shown to be TLR9-independent (Okabe *et al.*, 2005; Yoshida *et al.*, 2005), but was dependent upon IRF3 and IRF7 activation, although the TNF-α production was not (Okabe *et al.*, 2008).

The embryonic lethality of *DNase II* null mice was rescued by deficiency of the type I IFN receptor gene (IFNAR) (Yoshida *et al.*, 2005). *DNase II*$^{-/-}$*/IFNAR*$^{-/-}$ mice were viable but developed a chronic polyarthritis, which closely resembled human rheumatoid arthritis (Kawane *et al.*, 2006). At the onset of pathogenesis, levels of TNF-α were elevated in the bone marrow of these mice, which suggested that the

failure of macrophages to degrade apoptotic/erythroid precursor DNA caused them to produce TNF-α. This in turn would activate synovial cells to produce other cytokines and so drive the development of chronic polyarthritis (Kawane *et al.*, 2006; reviewed by Nagata *et al.*, 2010). Therefore, the consequence of loss-of-function mutations in *DNase II* provides a further example of the accumulation of unprocessed nucleic acids leading to disease.

1.3.3. RNase L and the IFN-induced antiviral pathway

The endoribonuclease RNase L is the key effector of the 2-5A system, an enzymatic RNA decay pathway activated by IFNs and dsRNA (Zhou *et al.*, 1993). RNase L is expressed in the majority of mammalian cell types but is latent until it is bound by small specialised oligonucleodides called 2′, 5′-oligoadenylates (2-5A), which are synthesised from ATP by 2-5A-oligoadenylate synthetase (OAS) proteins, in response to IFN (Rebouillat and Hovanessian, 1999; Zhou *et al.*, 1993). The OAS proteins are constitutively expressed at low levels and can act as pattern recognition receptors (PRRs) for the detection of viral dsRNA in the cytoplasm (reviewed by Sadler and Williams, 2008). Activated RNase L dimerises and cleaves ssRNA at UpN sequences, resulting in products with a 3′-phosphate (Floyd-Smith *et al.*, 1981). RNase L degrades viral RNA as a method of intrinsic cellular defence but can also cleave self RNAs (reviewed by Li *et al.*, 2010). RNA degraded by RNase L can act as a ligand for the cytoplasmic RNA helicases RIG-I and MDA5, leading to the induction of a type I IFN response (Malathi *et al.*, 2007). In addition to this role in the antiviral response, RNase L has been shown to be required for an effective antibacterial response (Li *et al.*, 2008).

A number of other cellular functions have been attributed to RNase L, including a role in myoblast differentiation (Bisbal *et al.*, 2000) and as a tumour suppressor (Al-Ahmadi *et al.*, 2009; Liu *et al.*, 2007; Zhou *et al.*, 1997). In addition to defects in immune signalling in response to pathogens, *RNASEL*-deficient mice have enlarged spleens and thymii as a consequence of a suppression of apoptosis (Zhou *et al.*, 1997).

Germline mutations in *RNASEL* have been identified in patients with hereditary prostate cancer (HPC) (Carpten *et al.*, 2002; Casey *et al.*, 2002). There is a particularly strong association of the *RNASEL* R462Q variant with the disease, with homozygotes having a two-fold increased risk of developing prostate cancer (Casey *et al.*, 2002). The mutation was shown to result in a three-fold decrease in the catalytic activity of the enzyme, due to impaired dimerisation of the protein (Casey *et al.*, 2002; Xiang *et al.*, 2003).

It was proposed that inherited defects in RNase L could permit infection with an oncogenic virus leading to disease. Subsequently, a novel retrovirus, xenotrophic endogenous murine leukaemia retrovirus (XMRV) was identified almost exclusively in HPC patients homozygous for reduced activity variants of the *RNASEL* gene (Dong *et al.*, 2007). XMRV cDNA was isolated from human prostate cancer tissue and shown to be infectious, via engagement with xenotrophic and polytrophic retrovirus receptor 1 (XPR1) (Dong *et al.*, 2007). The virus was sensitive to IFN, which strongly implicated RNase L in the IFN antiviral response against XMRV (Dong *et al.*, 2007; Fan, 2007). Interestingly, XMRV does not contain any host-derived oncogenes, nor was it shown to integrate within/near any known proto-oncogenes or tumour suppressor genes, therefore it has been hypothesised that its involvement in the pathogenesis of prostate cancer is via a paracrine mechanism (Kim *et al.*, 2008a; Summers and Crespi, 2008). Germline *RNASEL* mutations leading to an increased risk of disease have since been described in patients with other types of cancer (Madsen *et al.*, 2008), implying that RNase L defects may not be limited to prostate cancer.

1.3.4. FEN-1, inflammation and cancer

Loss-of-function mutations in the gene encoding the FEN-1 nuclease, which is involved in DNA replication and repair, have also been linked to human disease. This structure-specific nuclease, formerly known as DNase IV, has three separate nuclease activities: flap-specific endonuclease (Harrington and Lieber, 1994), 5′-

exonuclease (Kikuchi *et al.*, 2005) and gap-dependent endonuclease activities (Zheng *et al.*, 2005). Heterozygous somatic mutations in *FEN1* have been identified in human cancers, the majority of which abrogate both the 5′-exonuclease and gap-dependent endonuclease activities of the enzyme (Zheng *et al.*, 2007). These activities are essential for DNA repair and apoptotic DNA fragmentation. Consequently, transgenic mice harbouring one of the human *FEN1* mutations exhibited an increased sensitivity to DNA-damaging agents, an impaired ability to undergo apoptosis, and accumulated incompletely digested DNA fragments in apoptotic cells (Zheng *et al.*, 2007). The mice subsequently developed anti-nuclear antibodies (ANAs) and ICs, which deposited in the kidneys and lungs, and a chronic inflammatory response (Zheng *et al.*, 2007), which therefore provides a further example of the activation of autoimmunity and inflammation as a consequence of failure to clear apoptotic debris.

In addition to the autoimmune phenotype, these mice had a greatly increased risk of tumourigenesis, suggesting that FEN-1-deficiency leads to a higher risk of mutations in tumour-suppressor genes and oncogenes (Zheng *et al.*, 2007). Additionally, the chronic inflammation promoted cancer progression in these mice, via activation of the transcription factor NF-κB (Zheng *et al.*, 2007).

In summary, nucleases function across a wide range of cellular processes. Whilst OAS-RNase L functions in a PRR pathway, the other nucleases described in this section play key roles in the removal of nucleic acid-containing material from dead cells. Failure to clear cellular debris induces a destructive inflammatory response. Therefore, loss-of-function mutations in genes encoding nucleases lead to autoimmune disease via activation of the innate immune system by their unprocessed nucleic acid substrates.

1.4. Cellular sources of RNA:DNA hybrids

As discussed in Section 1.2, loss-of-function mutations in the three genes encoding the subunits of the RNase H2 enzyme complex cause AGS. Loss of RNase H2 activity would be expected to result in increased levels of its unprocessed substrates, as occurs in DNase I- and DNase II-deficient cells. The accumulation of TREX1 substrates in cells from AGS patients with loss-of-function mutations in *TREX1* has been reported (Yang *et al.*, 2007). The inappropriate activation of the innate immune system by such self-nucleic acids is hypothesised to be responsible for the development of the inflammatory response seen in cells lacking TREX1 activity (Stetson *et al.*, 2008).

Endogenous RNA:DNA hybrids, the substrates of RNase H enzymes, are naturally formed during several cellular processes. The accumulation of RNase H2 substrates in cells from AGS patients has not yet been published. Therefore, understanding the nature of endogenous RNA:DNA hybrids will assist in the identification of substrates which may be immunostimulatory when unprocessed.

1.4.1. DNA replication and Okazaki fragments

The successful replication of DNA during S phase of the cell cycle is an essential cellular process. The DNA is first unwound by helicases, forming a replication fork structure (Figure 1.9). The two strands of the fork, the leading strand and the lagging strand, serve as templates for replication. In eukaryotes, the DNA polymerase-α (pol-α)/primase complex is recruited to these structures where it initiates replication by synthesising short (8 - 12 nt long) RNA primers. The primase tightly binds these RNA primers to protect them from degradation (Yuzhakov *et al.*, 1999).

On the leading strand, the primase complex is displaced and the nascent DNA synthesised in a continuous manner by pol-ε (Lee *et al.*, 1991; Pursell *et al.*, 2007). However, as polymerases can only synthesis DNA unidirectionally (5′ → 3′), the lagging strand is replicated discontinuously, with small stretches of DNA synthesised at a time, which requires frequent re-priming (Figure 1.9).

The DNA polymerase-α/primase complex switches from primase activity to DNA polymerase activity and extends the RNA primers by synthesising short (approximately 20 nt) stretches of initiator DNA (Arezi and Kuchta, 2000; Waga and Stillman, 1994). The pol-α complex is then displaced by the higher fidelity DNA polymerase pol-δ, facilitated by the loading of PCNA (reviewed by Burgers, 2009; Tsurimoto and Stillman, 1991). Pol-δ then further elongates the chimeric RNA-DNA primer (Okazaki fragment) in a progressive manner until it encounters the 5′ end of the previously synthesised Okazaki fragment (reviewed by Waga and Stillman, 1998). The synthesis and joining of Okazaki fragments is a highly conserved process, found in all prokaryote and eukaryote cells (reviewed by Kao and Bambara, 2003).

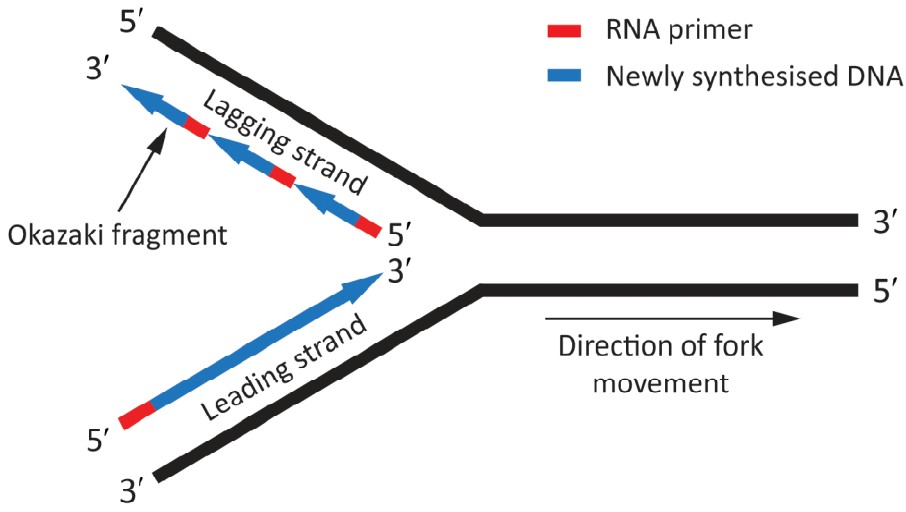


Figure 1.9. Okazaki fragments formed during lagging strand replication.
Replication of the lagging strand involves frequent repriming of the strand, resulting in many RNA-DNA initiator fragments.

Removal of the initiator RNA-DNA primers is required for DNA replication to be completed, as DNA ligase 1 will not efficiently ligate RNA-DNA nicks (Stith *et al.*, 2008). Pol-α lacks proof-reading activity, therefore part of the initiator DNA as well as the RNA has to be removed to maintain replication integrity (reviewed by Kao

and Bambara, 2003). Pol-δ is able to proof-read the nascent DNA for errors made by pol-α (Pavlov *et al.*, 2006). Upon encountering the downstream Okazaki fragment, the 5′-end of the primer is displaced, creating an RNA-containing single-stranded flap. The size of the flap is determined by the extent of elongation of the fragment by pol-δ (reviewed by Burgers, 2009; Stith *et al.*, 2008). Short flaps (1-2 nt) are cleaved by FEN-1 to generate a nick, which is filled in by DNA ligase 1 (Ayyagari *et al.*, 2003; Burgers, 2009). The flap can then undergo nick translation until all of the initiator RNA is removed (reviewed by Burgers, 2009). The maturation of Okazaki fragments in this way results in a complete strand of newly synthesised DNA. FEN-1 activity is stimulated by interaction with PCNA (Maga *et al.*, 2001) and by the activity of the WRN helicase in human cells (Brosh *et al.*, 2001; Chakraborty and Grosse, 2010). Longer flaps or flaps adopting secondary structures are processed by an additional multifunctional nuclease, DNA replication helicase 2 homologue (Dna2) (Bae *et al.*, 2001; Bae and Seo, 2000).

RNase H enzymes were implicated with a role in the removal of the short RNA primers, in cooperation with FEN-1 (Bae *et al.*, 2001; Bambara *et al.*, 1997; Turchi *et al.*, 1994). However, studies in *S.cerevisiae* suggest that RNase H2 is functionally redundant with Rad27p (the FEN-1 orthologue) (Ii and Brill, 2005) and instead indicate that Rad27p and Dna2 act to remove the RNA primers in an RNase H-independent manner (Bae and Seo, 2000; Rossi and Bambara, 2006).

In *Xenopus laevis*, it has been shown that new RNA-DNA primers continue to be synthesised by pol-α downstream of stalled replication forks, leading to an accumulation of these primers which activate Chk1-mediated cell cycle checkpoints (Van *et al.*, 2010). It has not yet been established if primer accumulation occurs in mammalian cells.

1.4.2. RNA:DNA hybrids and transcription

An R loop is formed when a newly transcribed RNA strand is base paired with one DNA strand of a DNA duplex, leaving the displaced DNA strand single-stranded

(Figure 1.10) (Thomas *et al.*, 1976). *In vivo*, stable R loops are known to exist in both prokaryotes and eukaryotes, arising as a consequence of the synthesis of a guanine (G)-rich RNA strand by an RNA polymerase transcribing a cytosine (C)-rich template (Lee and Clayton, 1996; Yu *et al.*, 2003). RNA:DNA hybrid duplexes consisting of a purine-rich RNA strand and a pryimidine-rich DNA strand have a greatly increased stability compared to the equivalent DNA duplex (discussed further in Chapter 4) and so the R loop structures persist preferentially in these regions (Ratmeyer *et al.*, 1994). It has recently been shown that a nick in the nontemplate strand can act as a site for R loop initiation, even in the absence of G-rich sequence (Roy *et al.*, 2010).

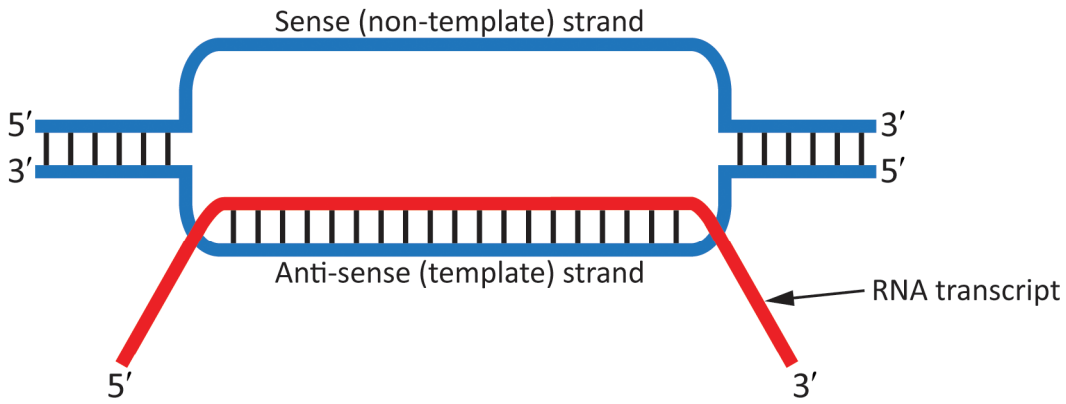


Figure 1.10. An R loop.
Schematic representation of the formation of an R loop. The G-rich newly-transcribed mRNA base pairs with the template strand, displacing the G-rich non-template strand.

1.4.2.1. R loops at origins of replication

The presence of relatively short (~140 bp) R loops at bacterial origins of replication has been well characterised (Baker and Kornberg, 1988; Itoh and Tomizawa, 1980; Masukata and Tomizawa, 1990). The transcript within the R loop is processed to

provide the RNA primer for synthesis of the leading strand (Belanger and Kreuzer, 1998) following unwinding of the R loop by helicase activity (Dudas and Kreuzer, 2001). The R loop therefore provides a method of regulating replication.

The mitochondrial origin of replication in eukaryote cells utilises a similar persistent R loop structure (Lee and Clayton, 1996; Xu and Clayton, 1995; Xu and Clayton, 1996). In addition to those priming leading strand replication at the origins of replication, supplementary R loops within the mitochondrial genome are proposed to have other roles, including altering the conformation of the mtDNA into a more open state (Brown *et al.*, 2008).

1.4.2.2. The impairment of transcription elongation by R loops

A key stage during the transcription of eukaryotic genomic DNA is the formation of a stable elongation complex comprising of RNA polymerase II (RNA pol II) and a minimal 8 bp R loop consisting of the template DNA strand hybridised to the newly transcribed mRNA strand (Kireeva *et al.*, 2000b). This hybrid has a key role in the stability and processivity of RNA pol II (Kireeva *et al.*, 2000a; Kireeva *et al.*, 2000b). However, under certain conditions, including the disruption of RNA binding proteins, much longer R loops can be formed co-transcriptionally behind the RNA pol II complex (Huertas and Aguilera, 2003). These R loops impair transcription elongation and lead to genetic instability by promoting transcription-associated recombination (Huertas and Aguilera, 2003; Lin *et al.*, 2010). Interestingly, this phenotype can be rescued in *S. cerevisiae* mutants by the overexpression of RNase H1, suggesting that the RNA:DNA hybrid needs to be removed for transcription elongation to proceed, presumably because they block the way for the next elongation by RNA pol II (Huertas and Aguilera, 2003). This provides evidence for a biological function for RNA:DNA hybrids, as a transcriptional intermediate able to reduce the efficiency of transcriptional elongation (Huertas and Aguilera, 2003; Tous and Aguilera, 2007).

The SR protein alternative splicing factor/splicing factor 2 (ASF/SF2) plays an important role in the suppression of R loop formation, by binding the nascent G-rich

RNA transcript (facilitated by RNA pol II), and consequently preventing the RNA from associating with the template DNA strand (Li and Manley, 2005). Eukaryote cells deficient in ASF/SF2 have extensive regions of R loops and subsequently develop DNA double-strand breaks (DSBs) which activate the cell cycle checkpoints and lead to arrest in G2 (Li and Manley, 2005; Li and Manley, 2006). Overexpression of RNase H1 suppresses this genomic instability phenotype (Li and Manley, 2005).

1.4.2.3. R loops and class switch recombination

The most extensively characterised persistent R loops in mammalian cells are formed following transcription of the highly repetitive G-rich switch regions, located directly upstream of each constant region in the genes encoding Immunoglobulin (Ig) in B cells. In response to activation by antigen, transcription is stimulated from specific promoters in order to switch the isotype of the constant region of the gene to a more specialised class, consequently switching Ig production from IgM to IgG, IgA or IgE (reviewed by Fugmann and Schatz, 2003). This class switch recombination (CSR) results in the formation of large (frequently > 1 kb) R loops in the switch regions of the gene (Daniels and Lieber, 1995; Reaban and Griffin, 1990; Reaban *et al.*, 1994; Yu *et al.*, 2003). Cytosines in the displaced single-stranded DNA are converted to uracils by the single-strand specific enzyme, Activation-Induced cytosine Deaminase (AID). Uracils are unable to pair with the corresponding guanines on the opposite strand, therefore the base excision repair machinery is activated. Subsequently DNA DSBs are induced within the switch regions. The two R loop regions, including the constant region between them, are excised and the remaining ends joined by the DNA repair complexes (Fugmann and Schatz, 2003; Huang *et al.*, 2006). RNA:DNA hybrid formation at these specific regions of the Ig genes is therefore an essential process in the function of the adaptive immune system.

1.4.3. RNA-templated DNA repair

The use of RNA as a template for DNA synthesis can occur in *S.cerevisiae* during the repair of DNA double strand breaks (DBSs) as the replicative DNA polymerases

α and δ are able to copy short RNA templates (Nick McElhinny *et al.*, 2010; Storici *et al.*, 2007). *In vitro* experiments showed that pol-α is able to copy four ribonucleotides embedded in a 45 nt partially double-stranded DNA molecule, (Storici *et al.*, 2007). This RNA-templated DNA repair generates a RNA:DNA hybrid as an intermediate, which could potentially block further elongation of the newly synthesised DNA by pol-δ (Storici *et al.*, 2007). It has been hypothesised that such RNA-templated DNA synthesis could occur in regions where RNA exists within DNA, such as the mitochondrial genome, or when large amounts of endogenous RNA are present, specifically during or immediately after transcription (Nick McElhinny *et al.*, 2010; Storici *et al.*, 2007).

1.4.4. RNA:DNA hybrids at telomeres

Telomeres consist of repetitive G-rich dsDNA bound to a multiprotein complex located at the ends of linear eukaryotic chromosomes, and play an important role in protecting chromosome ends from degradation (reviewed by de Lange, 2005). In eukaryotic cells, telomeric DNA repeats are transcribed into telomeric repeat-containing RNA (TERRA) (Azzalin *et al.*, 2007; Luke *et al.*, 2008; Schoeftner and Blasco, 2008) by RNA pol II (Horard and Gilson, 2008; Luke *et al.*, 2008). These non-coding TERRA molecules range in size from ~100 nt to ~9000 nt, and can form an RNA:DNA hybrid type G-quadruplex structure with the telomeric DNA (reviewed by Phan, 2010; Xu *et al.*, 2009). In yeast, it has been shown that a hybrid formed between TERRA and DNA inhibits the activity of telomerase, a specialised RNP enzyme complex responsible for telomere elongation (Luke *et al.*, 2008). This inhibition can be overcome *in vitro* by overexpression of RNase H, however resolution of these hybrids is thought to occur in an RNase H-independent mechanism *in vivo* (Luke *et al.*, 2008). Therefore, it is hypothesised that the primary function of such RNA transcripts is the maintenance of higher-order telomeric chromatin structures via the formation of a stable RNA:DNA hybrid at these regions which block telomerase binding, and lead to telomere shortening (Azzalin *et al.*, 2007; O'Sullivan and Karlseder, 2010).

Telomerase, which is essential during development but is switched off in mature cells (with the exception of cancer cells (Kim *et al.*, 1994)), is active as a high-fidelity reverse transcriptase, adding multiple identical repeats of DNA to the 3′-ends of telomeres (Shippen-Lentz and Blackburn, 1990). Replication is initiated following the pairing of the RNA-templating region, a small nucleolar RNA called telomere RNA component (TERC) with the incoming ssDNA primer to be extended (Shippen-Lentz and Blackburn, 1990). This forms an initial RNA:DNA hybrid region of 4 - 5 bp (Gilley *et al.*, 1995). Following one round of telomere replication, the hybrid is transiently dissociated before it repositions at the other end of the template. The binding of the protein component of telomerase, telomere reverse transcriptase (TERT), to the RNA:DNA hybrid is structurally very similar to the binding of retroviral HIV-1 reverse transcriptase (RT) to its substrate, supporting a common mechanism of DNA replication between the two enzyme families (Mitchell *et al.*, 2010).

1.4.5. Endogenous retroelements

A large proportion of mammalian genomes are derived from endogenous retroelement sequences (Deininger and Batzer, 2002). Retroelements are characterised by the ability to mobilise to another genomic location using an RNA intermediate. This RNA is reverse transcribed, forming a transient RNA:DNA hybrid, before conversion into dsDNA and integration into the genome (reviewed by Malik, 2005). A subcategory of retroelements is endogenous retroviruses (ERVs): ancient proviral remnants of retroviral infection of the germ-line. Most have undergone extensive deletions, however some have retained the ability to code for functional proteins (reviewed by Bannert and Kurth, 2004). Activation of human ERVs (HERVs) has been implicated as a causative trigger of autoimmune disease (Balada *et al.*, 2009).

1.4.6. Incorporation of ribonucleotides into DNA

It is not known to which extent ribonucleotides are incorporated into DNA in the place of deoxyribonucleotides during normal replication *in vivo*, particularly as the normal cellular environment contains both but an excess of ribonucleotides (Nick McElhinny *et al.*, 2010). In addition to the misincorporation of rNTPs by error-prone DNA polymerases described in Section 1.1.4.2., incidences of ribonucleotide incorporation into DNA have been described in replication intermediates purified from human mitochondria (Pohjoismaki *et al.*, 2010; Yang *et al.*, 2002b; Yasukawa *et al.*, 2006).

The active incorporation of ribonucleotides as an epigenetic mark has been described in fission yeast. The *Schizosaccharomyces pombe mat1* imprint which marks the cells that are able to undergo mating-type switching, consists of two adjacent ribonucleotides in one strand of the genomic DNA (Vengrova and Dalgaard, 2004; Vengrova and Dalgaard, 2006). Switching initiates when the leading-strand replication complex encounters the RNA imprint in the template strand. It is unlikely that the imprint is formed from an RNA primer remaining after the processing of Okazaki fragments and may instead be created by ribose oxidation (Vengrova and Dalgaard, 2004). It has not yet been confirmed if the ribonucleotides are comprised of uracil or ribothymine (Vengrova and Dalgaard, 2006), however, the ribonucleotide imprint provides an example of a naturally occurring RNase H2-specific substrate.

In summary, endogenous RNase H2 substrates arise as a consequence of various biological processes, and can play an essential role in numerous cellular processes. It is conceivable that RNase H2 deficiency could result in accumulation of any of these RNA:DNA hybrid-containing nucleic acid species, triggering the inflammatory disease state and subsequent pathology seen in AGS.

1.5. Nucleic acids as activators of the innate immune response

The innate immune response is the first-line host defence against pathogenic infection. It is an evolutionarily ancient system, found in both the plant and animal kingdoms, in contrast to the adaptive immune system which is found only in jawed vertebrates (Agrawal *et al.*, 1998; reviewed by Hoffmann *et al.*, 1999). Through the genetic rearrangement of receptor gene segments, the adaptive immune response can produce T and B lymphocytes targeted to specific pathogens, providing a highly specialised and effective response. The receptors of the innate immune system are fixed in the genome and do not require rearrangement, resulting in an immediate primary immune response to pathogenic invasion, and subsequent stimulation of adaptive immunity (reviewed by Iwasaki and Medzhitov, 2010).

1.5.1. Innate immune recognition of pathogens

The innate immune system is activated in response to the detection of pathogen-associated molecular patterns (PAMPs), present in pathogens but normally absent in host cells. The pattern recognition receptors (PRRs) of the mammalian innate immune system recognise a wide variety of PAMPs, including microbial structural components and nucleic acids early in the infection process.

PRR signalling can also be activated in the absence of infection by the release of endogenous danger-associated molecular patterns (DAMPs), usually from damaged cells (Matzinger, 1994). DAMPs can include proteins, such as HMGB1 (Klune *et al.*, 2008), and non-protein molecules, including DNA (Imaeda *et al.*, 2009; reviewed by Rubartelli and Lotze, 2007; Zhang *et al.*, 2010).

PRRs can be broadly divided into four classes, based on their cellular location (reviewed by Kawai and Akira, 2009). The Toll-like receptor (TLR) and C-type lectin-like receptor (CLR) families are transmembrane proteins, located at the plasma membrane or the endosomal membrane. The Retinoic acid-Inducible Gene I (RIG-I)-like receptors (RLRs) and Nucleotide-binding Oligomerisation Domain (NOD)-like

receptors (NLRs) are cytosolic receptor families. All of these receptors are widely expressed; however specialist cells of the innate immune system, including dendritic cells (DCs) and macrophages express specific subsets.

1.5.2. Innate immune recognition of nucleic acids

Nucleic acids are potent activators of the innate immune response (Table 1.2). The detection of viral and bacterial nucleic acids by intracellular PRRs rapidly triggers a series of signalling cascades which induce a powerful inflammatory response (reviewed by Yoneyama and Fujita, 2010). Activation of these PRRs is particularly characterised by the production of type I IFNs, which have key roles in the antiviral response, and proinflammatory molecules which drive the inflammatory response.

Distinguishing pathogenic 'non-self' nucleic acids from 'self' nucleic acids is crucial to prevent the inappropriate activation of innate immunity. As described in Section 1.3, the triggering of autoimmunity by nucleic acids is strongly linked with the pathogenesis of inflammatory disorders, including SLE, rheumatoid arthritis and AGS. Consequently, activation of innate immune responses to nucleic acids is tightly regulated. In contrast to many autoimmune diseases which have a complex genetic basis, AGS provides a monogenic model for understanding the role of nucleic acids in the innate immune response.

1.5.3. Membrane-bound sensors of nucleic acids

TLRs are single transmembrane proteins consisting of an N-terminal leucine rich-repeat (LRR) region and a cytosolic Toll/Interleukin-1 (IL-1) receptor (TIR) signal transduction domain (Rock *et al.*, 1998). They are functional as dimers, predominantly homodimers although heterodimers also occur. Ligand binding to the LRR induces a conformational change in the receptor which allow binding of the TIR to adapter proteins recruiting additional proteins, leading to a signal transduction cascade (reviewed by O'Neill and Bowie, 2007). Despite these similarities, each TLR is highly specialised at recognising specific PAMPs (Table 1.2).

Table 1.2. Recognition of nucleic acids by PRRs.

	Name	Location	Adapter	Pathogenic nucleic acids	Synthetic ligands	References
TLRs	TLR3	Endosome[a]	TRIF	Viral dsRNA	Poly (I:C), siRNA	Alexopoulou et al. 2001, Okahira et al. 2005, Kleinman et al. 2008
	TLR7	Endosome	MyD88	Viral ssRNA	Imidazoquinolines, U-rich ORNs	Hemmi et al. 2002, Jurk et al. 2002, Diebold et al. 2004, Heil et al. 2004, Lund et al. 2004, Diebold et al. 2006
	TLR8[b]	Endosome	MyD88	Viral ssRNA	Imidazoquinolines, GU-rich ORNs	Hemmi et al. 2002, Jurk et al. 2002, Heil et al. 2004
	TLR9	Endosome	MyD88	Bacterial DNA, viral DNA	CpG ODNs	Hemmi et al. 2000, Bauer et al. 2001, Lund et al. 2003, Krug et al. 2004,
RLRs	RIG-I	Cytosol	IPS-1	Viral RNA	IVT RNA, short 5'ppp dsRNA	Hornung et al. 2004, Pichlmair et al. 2004, Schlee et al. 2009, Schmidt et al. 2009, Rehwinkel et al. 2010
	MDA5	Cytosol	IPS-1	Viral RNA	Poly (I:C)	Gitlin et al. 2006, Kato et al. 2006, Pichlmair et al. 2009
	LGP2	Cytosol		Viral RNA		Murali et al. 2008, Pippig et al. 2009
RNA sensors	PKR	Cytosol		Viral RNA	Poly (I:C)	Meurs et al. 1990, Gil et al. 2004
NLRs	NLRP3 inflammasome	Cytosol	ASC	Bacterial DNA, viral DNA, viral RNA	Poly (dAdT), dsDNA (> 250bp), poly (I:C), ORNs	Muruve et al. 2008, Allen et al. 2009
	AIM2 inflammasome	Cytosol	ASC	Viral dsDNA	Poly (dAdT), dsDNA (> 100bp)	Burckstrummer et al. 2009, Fernandes-Alnemri et al. 2009, Hornung et al. 2009, Roberts et al. 2009
DNA sensors	DAI	Cytosol			Poly (dAdT), ISD	Takaoka et al. 2007

[a]TLR3 is predominantly expressed intracellularly but can be detected at the cell surface in some fibroblast cell lines and primary astrocytes (Bsibsi *et al.*, 2002). [b]Human TLR8 only

Within the human TLR family, there are four receptors for nucleic acids; TLR3, TLR7, TLR8 and TLR9. All are predominantly located on intracellular membranes; including endosome, lysosomes and the endoplasmic reticulum (Latz *et al.*, 2004; Nishiya and DeFranco, 2004; Nishiya *et al.*, 2005). The intracellular localisation of TLR3 is dependent on a cytoplasmic linker region between the TIR and transmembrane domains (Funami *et al.*, 2004; Nishiya *et al.*, 2005), whilst TLR7 and TLR9 localisation is mediated instead through the transmembrane domains (Barton *et al.*, 2006; Nishiya *et al.*, 2005). This suggests that their primary function could be the recognition of extracellular nucleic acids, including those derived from pathogens or those released from damaged cells (Kawai and Akira, 2006). Alternatively, nucleic acids produced by viral replication within the cytosol of a cell could be captured within membranous vesicles and brought to the endosomes (Sen and Sarkar, 2005). There is conservation between the nucleic acid TLRs of human and mice but the expression of the receptors on subtypes of cells differs (Rehli, 2002).

In resting cells, TLR3, TLR7 and TLR9 interact with the membrane protein Unc93b1 in the endoplasmic reticulum (ER) (Brinkmann *et al.*, 2007). Upon stimulation with their ligands, Unc93b1 delivers TLR7 and TLR9 to the endosomes where they interact with their ligands in the acidic environment (Kim *et al.*, 2008b). Mutations in *Unc93b1* abolish signalling through these TLRs (Tabeta *et al.*, 2006). TLR7 and TLR9 signalling is also dependent upon trafficking to the endosome by another ER protein, the chaperone PRAT4A (Takahashi *et al.*, 2007). The necessary relocalisation of TLR7 and TLR9 to an environment where they recognise their ligands provides an additional control step in the avoidance of inappropriate recognition of self nucleic acid.

Interestingly, in DCs Unc93b1 was found to differentially bind to TLR7 and TLR9, biasing the response to nucleic acid from RNA-sensing to DNA-sensing (Fukui *et al.*, 2009), which may have important implications for autoimmune disease (discussed further in Section 1.5.3.4) (Fukui *et al.*, 2009; McGettrick and O'Neill, 2010).

1.5.3.1. TLR3

TLR3 detects dsRNA and can also interact with the synthetic analogue of dsRNA, polyinosine-polycytidylic acid (poly (I:C)) (Alexopoulou *et al.*, 2001; Okahira *et al.*, 2005). dsRNA is a predominantly viral product, present in the genome of dsRNA viruses and as a replication intermediate of ssRNA, ssDNA and dsDNA viruses, therefore TLR3 plays a role in the response to infection by a variety of different viral species (Tabeta *et al.*, 2004; Wang *et al.*, 2004; Zhang *et al.*, 2007). TLR3 can also recognise short interfering RNA (siRNA) in a sequence-independent manner (Kleinman *et al.*, 2008).

TLR3 mRNA is widely expressed across many tissues and most immune cells (Applequist *et al.*, 2002; Rock *et al.*, 1998). The TLR3 protein is expressed intracellularly by most cells, including conventional/myeloid DCs (cDCs) (Matsumoto *et al.*, 2003) but can be detected on the cell surface of some fibroblast cell lines (Matsumoto *et al.*, 2002). Plasmacytoid DCs (pDCs) do not express TLR3 (Hornung *et al.*, 2002; Kadowaki *et al.*, 2001; Okada *et al.*, 2003).

Signalling through TLR3 requires the adapter protein TIR-domain-containing adapter-inducing interferon-β (TRIF) (Oshiumi *et al.*, 2003). TRIF recruits the signalling molecules Receptor-Interacting Protein 1 (RIP1) and Tumour Necrosis Factor (TNF)-Associated Factor 6 (TRAF6). This leads to the subsequent activation of TRAF3 and Transforming Growth Factor β-Activated Kinase (TAK1) and their downstream signalling pathways. Subsequently, expression of proinflammatory cytokines is induced (by the NF-κB transcription factor complex) and type I IFNs (by IRF3 and IRF7) (reviewed by Kawai and Akira, 2009; and Yoneyama and Fujita, 2010) (Figure 1.11).

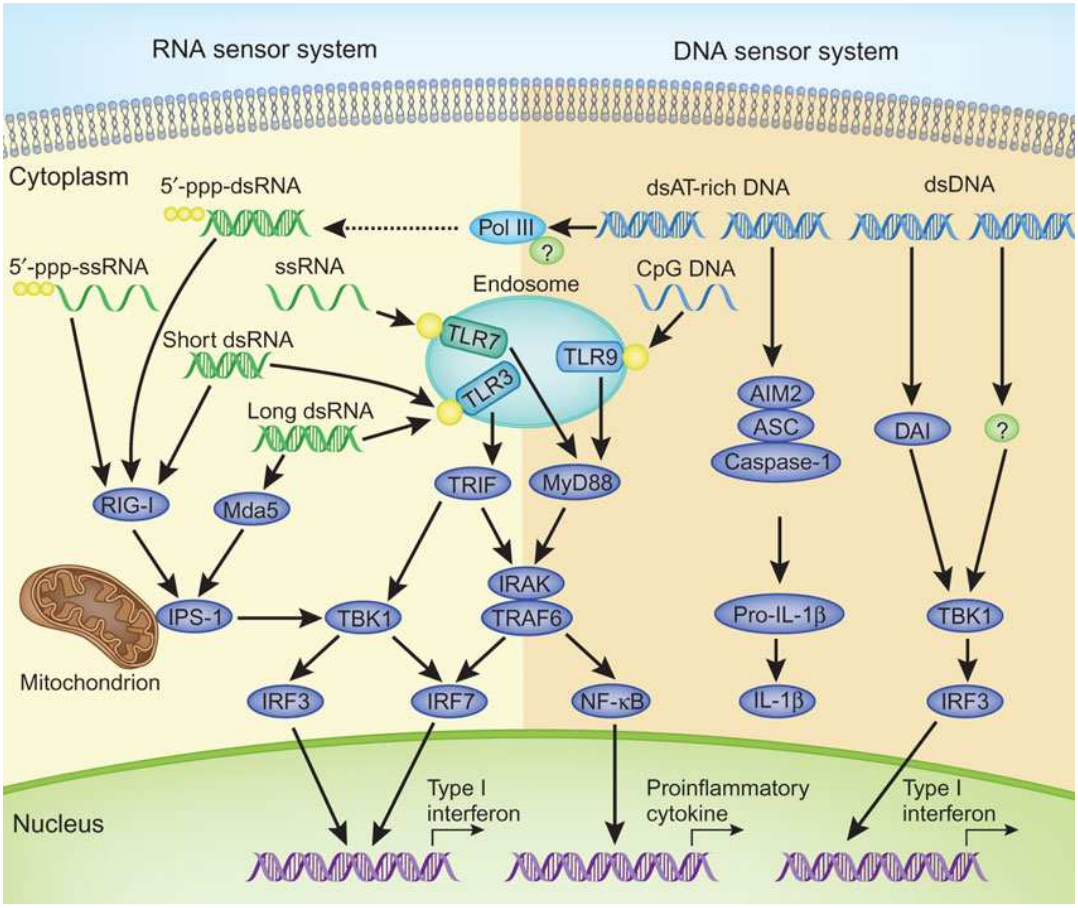


Figure 1.11. Activation of the innate immune system by nucleic acids.
Intracellular PRRs, including TLRs, RLRs, NLRs and the cytosolic DNA sensors DAI and RNA pol III detect pathogenic and self nucleic acids, triggering downstream signalling cascades (simplified in this diagram). IRF3 and IRF7 induce type I IFN production, whilst NFκB stimulates the transcription of proinflammatory cytokines. Both DNA and RNA can induce production of type I IFNs. Figure reproduced from Cao *et al*. 2009.

1.5.3.2. TLR7 and TLR8

TLR7 and TLR8 recognise viral ssRNA, and TLR7 is required for a type I IFN response against ssRNA viruses (Diebold *et al.*, 2004; Lund *et al.*, 2004). *In vitro* stimulations using synthetic oligoribonucleotides (ORNs) and purified viral RNA have further refined the nature of the TLR7 nucleic acid ligand. ssRNA containing a short stretch of uridines is necessary and sufficient for TLR7 activation in murine pDCs, when targeted to the endosome using liposomal transfection reagents (Diebold *et al.*, 2004; Diebold *et al.*, 2006; Heil *et al.*, 2004; Lund *et al.*, 2004). Synthetic

ligands for TLR7 include imidazoquinoline derivatives, which are small anti-viral compounds (Hemmi *et al.*, 2002; Jurk *et al.*, 2002).

In humans, the function of TLR7 overlaps with the closely related TLR8 protein (Du *et al.*, 2000). The imidazoquinoline resiquimod R848 can activate human TLR7/8 and murine TLR7 but not murine TLR8 (Hemmi *et al.*, 2002). In contrast the function of murine TLR8 (originally thought to be non-functional (Jurk *et al.*, 2002)), is unclear, as murine TLR7 responds to ligands that are specific for human TLR8 (Gorden *et al.*, 2006; Heil *et al.*, 2004). It has been recently suggested that TLR8 is required for the activation of murine pDCs in response to AT-rich viral DNA (Martinez *et al.*, 2010).

As with other TLRs, TLR7 expression is heterogenous between different immune cell types (Hornung *et al.*, 2002; Kadowaki *et al.*, 2001; Reis e Sousa, 2004). It is highly expressed by both human and murine pDCs and B cells (Edwards *et al.*, 2003; Gururajan *et al.*, 2007; Hornung *et al.*, 2002; Kadowaki *et al.*, 2001).

Signalling through TLR7 is dependent on its association with the adapter protein Myeloid Differentiation primary response gene (88) (MyD88) (Hemmi *et al.*, 2002) and subsequent recruitment of the signalling molecules Interleukin-1 Receptor-Associated Kinase 1 (IRAK1), IRAK4 and TRAF6 (Figure 1.11). Downstream signalling through IRF3/7 induces cytokine expression. In both mouse and human, signalling through TLR7 leads to the production of large amounts of type I IFN by pDCs and the production of pro-inflammatory cytokines, including IL-6, IL-12 and TNF-α, by cDCs (Heil *et al.*, 2004; Takaoka *et al.*, 2005). In pDCs, the recognition of viral ssRNA within the endosome by TLR7 is independent of infection by live virus or viral replication (Beignon *et al.*, 2005; Lund *et al.*, 2004). In addition, pDCs also utlise TLR7 for the detection of cytosolic viral replication intermediates that are transported into lysosomes via autophagy (Lee *et al.*, 2007).

1.5.3.3. TLR9

The ligand for TLR9 is ssDNA; it was originally thought to respond exclusively to bacterial DNA, which contains non-methylated CpG dinucleotides rarely found in vertebrate DNA (Bauer *et al.*, 2001; Hemmi *et al.*, 2000). Studies using DNA oligonucleotides (ODNs) with a phosphorothiorate (PS)-modified backbone (to improve the nuclease resistance of the ODN) further defined the TLR9 ligand as a nonmethylated CpG flanked 5′ by two purine bases and 3′ by two pyrimidines (Krieg, 2002). However, it has subsequently been shown that TLR9 detects unmodified ssDNA in the endosome in a CpG-independent manner, due to the immunostimulatory ability of the natural phosphodiester (PD) backbone, compared to the immunoinhibitory effect of a PS-modified backbone (Haas *et al.*, 2008). pDCs can produce large amounts of IFN-α as a consequence of TLR9 stimulation by both viral (Krug *et al.*, 2004; Lund *et al.*, 2003) and mammalian DNA (Lande *et al.*, 2007). In all cases, a TLR9 response is critically dependent on the compartmentalisation of the DNA into the endosomes (Guiducci *et al.*, 2006).

The expression of TLR9 closely resembles that of TLR7, with the highest expression found in pDCs and B cells (Bourke *et al.*, 2003; Hoene *et al.*, 2006; Hornung *et al.*, 2002). Also, as with TLR7-mediated signalling, MyD88 is required as the adapter protein for TLR9 (Hemmi *et al.*, 2000) (Figure 1.11).

As described in Section 1.5.2.1, TLR9 is located in the ER of resting cells but relocates to endosomes to engage with its ligand (Latz *et al.*, 2004; Leifer *et al.*, 2004). Upon entry into the endosome, the ectodomain is cleaved by cathepsins into a shorter processed form (Ewald *et al.*, 2008; Park *et al.*, 2008). Signal transduction occurs only through the cleaved form of TLR9, suggesting that the need for cleavage of receptor adds another step of regulation of an immune response through this receptor. It is not yet clear if similar proteolytic cleavage is required to generate a functional form of TLR7 (Ewald *et al.*, 2008; Park *et al.*, 2008).

1.5.3.4. Intracellular TLRs and autoinflammatory disorders

The intracellular TLRs are strongly linked to disease pathogenesis in a number of chronic inflammatory diseases, including SLE (reviewed by Drexler and Foxwell, 2010). The loss of tolerance to self-antigens, particularly nucleic acids, and subsequent development of auto-antibodies against them is a key stage in the development of a disease state in SLE (Barrat *et al.*, 2005; Ehlers and Ravetch, 2007). The pre-existing circulating auto-antibodies in the sera of SLE patients bind self nucleic acid, forming ICs. As discussed in Section 1.3, a potential source of self-nucleic acid is cellular material released from dying cells and a failure to clear this leads to the development of an autoimmune pathology.

ICs containing self-RNA and self-DNA facilitate detection of nucleic acid by B cells and pDCs via TLR7 and TLR9 (Lau *et al.*, 2005; Leadbetter *et al.*, 2002; Vollmer *et al.*, 2005; Yasuda *et al.*, 2007). Activation of B cells via association with the B cell receptor (BCR), leads to the production of class-switched anti-DNA and anti-RNA auto-antibodies. The production of autoantibodies in response to autoantigens engaging with and activating the TLRs is referred to as the Toll hypothesis (Martin and Elkon, 2005).

Activation of pDCs stimulates the production of vast amounts of cytokines and chemokines. The presentation of antigen to autoreactive T cells by pDCs further amplifies the involvement of adaptive immunity in SLE. IFN-α, produced in high levels by stimulated pDCs, has an immunoregulatory role in maintaining a positive feedback loop for immune activation. It exerts multiple effects on other immune cells, inducing the maturation of immature DC subsets and the activation of natural killer (NK) cells, T cells and B cells (Le Bon *et al.*, 2001; Nguyen *et al.*, 2000). This in turn drives auto-antibody secretion and IC formation. In maintaining the activation of immunity, IFN-α stimulates the inflammatory response which can lead to tissue damage.

Expression of TLR7 and TLR9 was found to be increased in peripheral blood mononuclear cells from SLE patients, which correlated with increased IFN-α

expression (Komatsuda *et al.*, 2008). In addition, expression of both the chaperone protein UNC93B1 and the adapter MyD88 have been shown to be increased in patients with active SLE which further implicated TLR7 and TLR9 in the pathogenesis of SLE (Nakano *et al.*, 2010). The investigation of murine models of SLE also revealed an association between increased levels of TLR7 and TLR9 and disease (Christensen *et al.*, 2006; Fairhurst *et al.*, 2008).

1.5.4. Cytosolic sensors of nucleic acids

There are a several well characterised cytosolic sensors of nucleic acids in mammalian cells, including members of the RLR and NLR families of receptors.

1.5.4.1. The role of RLRs in the recognition of cytosolic RNA

The RLR family contains three members; RIG-I, Melanoma Differentiation-Associated gene 5 (MDA5) and Laboratory of Genetics and Physiology 2 (LGP2) (Kang *et al.*, 2002; Rothenfusser *et al.*, 2005; Yoneyama *et al.*, 2004). They all share a central DExD/H-box RNA helicase domain and a C-terminal regulatory domain (CTD). RIG-I and MDA5 also contain two N-terminal caspase activation and recruitment domains (CARD) but LPG2 does not (Rothenfusser *et al.*, 2005). The IFN response induced by ligand binding to RIG-I and MDA5 is mediated through the adapter Interferon-β Promoter Stimulator 1 (IPS-1), which is a mitochondrial membrane protein (Kawai *et al.*, 2005; Meylan *et al.*, 2005; Seth *et al.*, 2005; Xu *et al.*, 2005). Downstream signalling promotes NF-κB-induced expression of proinflammatory cytokines and IRF3/7-induced expression of type I IFNs (Kumar *et al.*, 2006) (Figure 1.11).

RLRs are essential for a robust IFN response following viral infection in some cell types, including fibroblasts and cDCs, but are non-essential in pDCs which utilise the TLR system instead (Kato *et al.*, 2005; Melchjorsen *et al.*, 2005)

1.5.4.1.1. RIG-I

The most extensively characterised RLR is RIG-I. The specific ligand for RIG-I was originally identified as ssRNA with a 5′-triphosphate (5′-ppp) using *in vitro* transcribed (IVT) ssRNA (Hornung *et al.*, 2006; Pichlmair *et al.*, 2006). Subsequent studies using synthetic 5′-ppp ssRNA have shown that optimal IFN-production via RIG-I requires the recognition of RNA sequences with short (minimum of 5 – 10 bases) base-paired regions including a blunt 5′-ppp end (Schlee *et al.*, 2009b; Schmidt *et al.*, 2009). The major natural ligand for RIG-I was identified as the genome of RNA viruses. The regions of RNA secondary structure found ssRNA virus genomes is sufficient to activate IFN-signalling when a 5′-ppp is present, independent of viral replication (Rehwinkel *et al.*, 2010; Schlee *et al.*, 2009b). Cytosolic RNA is abundant in mammalian cells, however the 5′-ppp is generally removed by cleavage or capping before the RNA leaves the nucleus, which provides a method of selectively distinguishing self RNA from viral RNA (reviewed by Yoneyama and Fujita, 2010). It is not yet clear if RIG-I exhibits sequence preferences towards its ligands (Saito *et al.*, 2008; Schlee *et al.*, 2009a; Schmidt *et al.*, 2009), but a 5′-ppp-cytosine has been shown to be the weakest stimulator of IFN production (Schlee *et al.*, 2009b).

Anti-viral signalling by RIG-I requires a two step process of ligand detection and binding. Initially, the 5′-ppp is verified as a PAMP and is strongly bound by the CTD (Cui *et al.*, 2008; Takahasi *et al.*, 2008; Wang *et al.*, 2010a). RIG-I then translocates along one strand of the base-paired RNA using its ATPase activity, without unwinding the RNA (Myong *et al.*, 2009). In the absence of a 5′-ppp, the translocase activity of RIG-I is suppressed by the CARDs (Myong *et al.*, 2009).

It is not clear whether additional substrates stimulate an type I IFN response via RIG-I, as results using short (< 1 kb) poly (I:C) and short dsRNA lacking a 5′-ppp have given variable results (Hausmann *et al.*, 2008; Kato *et al.*, 2008; Saito and Gale, 2008; Schlee *et al.*, 2009a). However, it appears that RIG-I can sense RNA cleavage products, generated from self RNA by activated RNase L, which although often containing base-paired regions, have 3′-monophosphate ends (Malathi *et al.*, 2007).

Therefore, the detection of sub-optimal ligands by RIG-I may still be sufficient to induce an IFN response.

In response to certain RNA viruses, RIG-I can signal independently of IPS-1 to produce the pro-inflammatory cytokine IL-1β. It associates with the inflammasome adapter Apoptotic Speck-like protein with a Caspase-activating recruiting domain (ASC), triggering caspase-1-dependent inflammasome formation in a manner similar to the dsDNA-induced AIM2 inflammasome (Poeck *et al.*, 2010) (Section 1.5.4.3).

1.5.4.1.2. MDA5

Although structurally very similar, RIG-I and MDA5 respond differentially to specific RNA viruses (Gitlin *et al.*, 2006; Kato *et al.*, 2006; reviewed by Rehwinkel and Reis e Sousa, 2010). Viral activation of MDA5 *in vitro* is dependent on the presence of high molecular weight viral RNA, which forms a branched structure containing both single- and double-stranded regions (Pichlmair *et al.*, 2009). Similarly, MDA5 responds to long (> 1 kb) poly (I:C) but RIG-I does not (Kato *et al.*, 2008; Kato *et al.*, 2006). This could be because long poly (I:C) also forms the higher-order structure specific for MDA5 recognition (Pichlmair *et al.*, 2009; Rehwinkel and Reis e Sousa, 2010).

1.5.4.1.3. LGP2

LGP2 binds both single- and double-stranded RNA via its CTD (Murali *et al.*, 2008; Pippig *et al.*, 2009). It was originally characterised as a negative regulator of the RIG-I-dependent anti-viral response (Rothenfusser *et al.*, 2005) but *in vivo* studies have since revealed that LGP2 actually functions as a positive regulator of RIG-I- and MDA5-mediated anti-viral immunity, in cDCs and MEFs but not pDCs (Satoh *et al.*, 2010). LGP2 was shown to act upstream of RIG-I and MDA5, and is hypothesised to facilitate binding of these helicases to viral RNA by unwinding complex structures or removing proteins from the viral ribonucleoprotein complexes (RNPs) to free the RNA genome (Satoh *et al.*, 2010). Accordingly, LGP2 was not required for the response to synthetic RNA including poly (I:C) and IVT 5′-ppp

dsRNA and ssRNA (Satoh *et al.*, 2010). However, LGP2 can also interact with IPS-1, potentially blocking its associations with downstream kinases and subsequent stimulation of IFN production (Komuro and Horvath, 2006). Its roles in the RIG-I/MDA5 pathways have not been fully elucidated.

1.5.4.1.4. Regulation of RLR signalling

In addition to control by LGP2, RIG-I-mediated anti-viral signalling can also be positively regulated by STimulator of IFN Genes (STING), a mitochondrial transmembrane protein which potently induces IRF3 (Ishikawa and Barber, 2008; Ishikawa *et al.*, 2009; Zhong *et al.*, 2008; Zhong *et al.*, 2009). STING forms a complex with IPS-1 and RIG-I, to which it recruits TRAF family member-Associated NF-κB activator (TANK)-Binding Kinase 1 (TBK1), which activates the downstream transcription factors, consequently stimulating type I IFN expression (Zeng and Chen, 2008). This appears to be cell type specific, as DCs and macrophages from STING-deficient mice responded normally to RNA virus infection whereas viral replication in fibroblasts lacking STING was increased (Ishikawa and Barber, 2008). The recently identified WD-repeat containing protein 5 (WDR5) may also play a key role in the assembly of a virally-induced signalling complex with IPS-1 and RIG-I at the mitochondrial membrane (Wang *et al.*, 2010b).

1.5.4.2. PKR

The IFN-regulated, dsRNA-dependent Protein serine/threonine Kinase R (PKR) has also been implicated in the production of IFNs following viral infection (Meurs *et al.*, 1990). Inactive PKR binds dsRNA in the cytosol and subsequently becomes activated via autophosphorylation events. It then phosphorylates the α subunit of eukaryotic protein synthesis Initiation Factor 2 (eIF2α), resulting in a global inhibition of protein synthesis in virally infected cells (reviewed by Cole, 2007). PKR requires a minimum of 33 bp dsRNA for activation (Zheng and Bevilacqua, 2004), however it has been proposed activation can occur following the binding of PKR to 5′-ppp ssRNA with limited secondary structure (Nallagatla *et al.*, 2007). The role of PKR in generating an anti-viral IFN response differs from the RLRs, which induce IFN gene expression. Instead, PKR acts post-transcriptionally, influencing

protein translation by regulating the stability of the IFN mRNA transcripts (Schulz *et al.*, 2010). The involvement of PKR in an antiviral IFN response is both virus (Schulz *et al.*, 2010) and cell type specific (Diebold *et al.*, 2003; Hornung *et al.*, 2004).

1.5.4.3. NLRs and nucleic acids

The NLR family of cytosolic proteins includes the Nucleotide-binding Oligomerization Domain-containing proteins NOD1 and NOD2, which have an established role in the response to intracellular bacterial components (reviewed by Ting *et al.*, 2010). In murine cells, *Nod2* expression is induced in response to transfection with 20 nt ssRNA or infection with a ssRNA virus (Sabbah *et al.*, 2009). The protein was shown to positively regulate type I IFN production via an interaction with IPS-1, although it is not yet clear if Nod2 directly interacts with viral ssRNA or if it forms a complex with RIG-I and IPS-1 (Sabbah *et al.*, 2009). Additionally, human NOD2 has been shown to interact with one of the OAS proteins (OAS2), which positively regulated RNase L activity, implying that it may also play a role in enhancing antiviral responses (Dugan *et al.*, 2009).

The pro-inflammatory cytokine response to viral infection is mediated by a member of another NLR subfamily. NLRP3 is triggered by a range of pathogen- and host-derived molecules, including nucleic acids, to form an inflammasome, via the adapter protein ASC (Allen *et al.*, 2009; Delaloye *et al.*, 2009; Muruve *et al.*, 2008). The NLRP3-inflammasome triggers the release of IL-1β and IL-18 through caspase-1-mediated cleavage of their precursors. Gain of function mutations in NALP3 are present in several autoinflammatory disorders (reviewed by Franchi *et al.*, 2009).

Inflammasome formation in response to cytosolic nucleic acids also occurs via the Haematopoietic Interferon-inducible Nuclear protein (HIN)-200 family member Absent In Melanoma 2 (AIM2). In response to direct binding of dsDNA, AIM2 forms an inflammasome with ASC and caspase-1, leading to the production of IL-1β (Burckstummer *et al.*, 2009; Fernandes-Alnemri *et al.*, 2009; Hornung *et al.*, 2009;

Roberts *et al.*, 2009). However, neither the NLRP3- nor the AIM2-inflammasome pathways are essential for IFN production in response to cytosolic dsDNA (Hornung *et al.*, 2009; Muruve *et al.*, 2008).

1.5.4.4. DAI

Transfection of dsDNA into cells is known to induce a TLR9-independent IFN response (Ishii *et al.*, 2006; Stetson and Medzhitov, 2006; Yasuda *et al.*, 2005b). The IFN-inducible protein DNA-dependent Activator of IRFs (DAI) was identified as a cytosolic sensor for dsDNA, increasing in expression in response to the synthetic B-form DNA poly(dA:dT) (Takaoka *et al.*, 2007). Poly (dA:dT) is also an activator of the AIM2 inflammasome, but only induces an IFN-response via DAI (Takaoka *et al.*, 2007). DAI responds to the introduction of bacterial, viral and mammalian DNA to the cytosol, inducing cytokine expression via activation of both the IRF3 and NF-κB pathways (DeFilippis *et al.*, 2010; Kaiser *et al.*, 2008; Stetson and Medzhitov, 2006; Takaoka and Taniguchi, 2008; Takaoka *et al.*, 2007). It does not interact with RNA *in vitro* (Takaoka and Taniguchi, 2008). The stimulation of DAI by cytosolic DNA is sequence-independent but is dependent on DNA length, with longer (in excess of 500 bp) dsDNA inducing a greater IFN response than short dsDNA (Takaoka and Taniguchi, 2008). The role of DAI as a PRR appears to be cell type specific (Lippmann *et al.*, 2008). Additionally, DAI-deficient mice still produce type I IFN in response to poly(dA:dT), suggesting the existence of other receptors capable of inducing an IFN response to this ligand (Ishii *et al.*, 2008). It is not yet clear how DAI distinguishes between pathogenic and host DNA *in vivo* but it is possible that nuclease digestion of self-DNA, which could act as a ligand, is required to prevent inappropriate activation of immunity via DAI (Takaoka and Taniguchi, 2008).

1.5.4.5. RNA polymerase III

RNA polymerase III (RNA pol III) is predominantly located in the nucleus, but also localises to the cytosol where it has been shown to act as a sensor for pathogenic dsDNA (Ablasser *et al.*, 2009; Chiu *et al.*, 2009). RNA pol III detects stretches of AT-rich dsDNA in the cytosol and uses it as a template for transcription. This generates 5′-ppp AU-rich dsRNA which acts as a ligand for RIG-I and stimulates

type I IFN induction (Ablasser *et al.*, 2009; Chiu *et al.*, 2009). The RNA pol III-dependent IFN induction pathway was demonstrated for synthetic dsDNA in the form of poly (dA:dT) but was also shown to be crucial for the innate immune response to some bacterial DNA and dsDNA viruses (Ablasser *et al.*, 2009; Chiu *et al.*, 2009) which were previously known to stimulate RIG-I (Choi *et al.*, 2009; Rasmussen *et al.*, 2009). Therefore the uncovering of RNA pol III as the primary sensor for poly (dA:dT) reaffirmed the specificity of the RLRs for sensing RNA.

1.5.4.6. Additional sensors of cytosolic DNA

Recently, two proteomic screens of DNA binding proteins identified novel sensors of cytosolic DNA. IFI16, which contains a pyrin domain and two HIN domains, directly binds to transfected dsDNA containing a viral sequence motif, leading to the recruitment of STING and IRF3/NF-κB mediated expression of type I IFNs and other innate immune response genes (Unterholzner *et al.*, 2010). Similarly, two members of the aspartate-glutamate-any amino acid-aspartate/histidine (DExD/H)-box helicase (DHX) family were identified as cytosolic sensors of CpG ODNs in human pDC cell lines (Kim *et al.*, 2010). DHX36 binds to type A CpG ODNs to trigger production of type I IFNs and other cytokines via IRF7 signalling, whereas the closely related protein DHX9 binds to type B CpG ODNs and initiates innate immune signalling via NF-κB-dependent pathways. Notably, both helicases require interactions with MyD88 to elicit downstream signalling, demonstrating a vital role for the adaptor protein in both endosomal and cytosolic DNA sensing pathways in pDCs (Kim *et al.*, 2010).

1.5.5.The production of type I IFNs in response to stimulation of the innate immune response by nucleic acids

As mentioned throughout this section, the triggering of the innate immune response by PRR detection of nucleic acids, both pathogenic and self, leads to the production of large amounts of type I IFN. The IFNs are a group of cytokines with numerous functions, and exhibit antiviral and antiproliferative actions, whilst modulating both the innate and adaptive immune responses. In humans, there are three major classes

of IFNs; type I IFNs, consisting of multiple IFN-α isoforms and single β, ε, κ and ω isoforms; type II IFNs consisting of IFN-γ; and type III IFNs consisting of the three isoforms of IFN-λ (reviewed by Bogdan, 2000; reviewed by Bogdan *et al.*, 2004; Kotenko *et al.*, 2003; Sheppard *et al.*, 2003).

The two major groups of type I IFN isoforms, IFN-α and IFN-β, signal through a common cell surface receptor complex called the type I IFN Receptor (IFNAR). This comprises two subunits (IFNAR1 and IFNAR2) that are both required to assemble a functional receptor complex (Mogensen *et al.*, 1999). The interaction of IFN with IFNAR activates a number of signal transduction pathways, the best characterised of which is the JAK-STAT pathway (reviewed by Taniguchi and Takaoka, 2002). The tyrosine kinases Tyrosine Kinase 2 (Tyk2) and Janus Kinase 1 (Jak1) are associated with the cytoplasmic tails of the two IFNAR subunits. Receptor binding by exogenous IFN-α/β activates the kinases and subsequently results in the phosphorylation of the signal transducer and activator (STAT) proteins. This induces assembly of the STATs into transcriptional activator complexes; including the IFN-α-Activated Factor (AAF) and IFN-Stimulated Gene Factor 3 (ISGF3) complexes. Upon translocation to the nucleus, they bind with high affinity to IFN-Stimulated Response Elements (ISREs) and IFN-γ Activated Sites (GASs) in the promoters of target genes. This leads to induction of the expression of numerous genes, collectively known as IFN-Stimulated Genes (ISGs), including IRFs, cytokines, chemokines and apoptosis-related genes (de Veer *et al.*, 2001; Samarajiwa *et al.*, 2009). Members of the IRF family of transcription factors, in particular IRF3 and IRF7, bind to the promoters of the IFN-α/β genes and induce expression, contributing to the feed-forward loop. Although IFN-α production is primarily controlled at the transcriptional level, there is variation in the affinity of the different IFN-α proteins and IFN-β for the type I IFN receptor. This affects the stability of the IFN-IFNAR complex and consequently the resultant downstream signalling (Cutrone and Langer, 1997; Jaks *et al.*, 2007).

1.5.5.1. Type I IFN production by plasmacytoid dendritic cells

Expression of type I IFNs can be rapidly and transiently induced by several factors, but they are produced most notably in response to pathogenic nucleic acids. Virtually all cells can produce type I IFNs in response to pathogens, but cells of the immune system, pDCs in particular, are by far the most potent producers of IFN-α (Siegal *et al.*, 1999).

pDCs are specialised for rapidly producing significant amounts of IFN-α immediately following viral infection, via TLR7 and TLR9 – the only TLRs they express. They constitutively express high levels of IRF7 mRNA (Coccia *et al.*, 2004; Izaguirre *et al.*, 2003; Kerkmann *et al.*, 2003) and initially produce IFN-α independent of the IFNAR-dependent feedback loop (Di Domizio *et al.*, 2009; Kerkmann *et al.*, 2003). IFN-α production is subsequently amplified via the feedback loop, which results in the sustained production of IFN (Kerkmann *et al.*, 2003). pDCs produce large amounts of IFN-α in response to TLR9 ligands (CpG-ODNs) and some TLR7 ligands (including the synthetic dsRNA poly (A:U), which is approximately 80 – 350 bp in size (Sugiyama *et al.*, 2008)).

Studies using the TLR9 ligands have revealed that pDCs have a unique mechanism of ligand transfer compared to cDCs, which further enhances their ability to mount a robust IFN response. In cDCs, CpG-ODN is rapidly transferred from the early endosome to the late endosome, stimulating the expression of pro-inflammatory cytokines and the co-stimulatory molecules CD40, CD80 and CD86 via NF-κB-dependent transcription. In contrast, CpG-ODN is retained in the early endosome of pDCs for much longer, allowing TLR9-MyD88 complexes to form and activated IRF7, resulting in a prolonged IFN response (Honda *et al.*, 2005). This mechanism is dependent upon the multimeric nature of Type A CpG-ODNs, which are capable of forming tertiary structures (Haas *et al.*, 2009; Honda *et al.*, 2005). In contrast, linear Type B CpG-ODNs only weakly stimulate IFN-α production by pDCs (Krieg, 2002). The retaining of TLR9 ligands in the endosome has also been shown to occur *in vivo*. In individuals with the autoimmune disorder of the skin psoriasis, self-DNA forms a complex with the antimicrobial peptide LL37, which has been shown to facilitate

uptake of the DNA by pDCs (Lande *et al.*, 2007; Nestle *et al.*, 2005). The DNA was shown to be retained in the early endosomes, where it stimulates a sustained IFN-α response via TLR9 (Lande *et al.*, 2007) and so drives the development of disease (Nestle *et al.*, 2005). This process may also occur in SLE patients, via engagement of the pDC FcγRIIa receptor by nucleic acid-containing ICs (as discussed in Section 1.3.1).

1.5.5.2. Differential expression of type I IFNs

The boundaries of the type I IFN gene locus, located on human chromosome 9 and murine chromosome 4, are defined by the evolutionarily conserved genes encoding IFN-β and IFN-ε (Hardy *et al.*, 2004). The number of genes encoding IFN-α family members varies between species: 13 protein-coding genes have been identified in human and 14 in mouse. All human and mouse IFN-α and IFN-β genes are intronless. The human IFN-α subtypes share approximately 70% amino acid sequence identity, and the IFN-α consensus sequence shares 35% identity with human IFN-β (Genin *et al.*, 2009).

The IFN-α subtypes are differentially expressed in response to viral infection, although this varies greatly depending on both the infecting virus and cell type (reviewed by Honda *et al.*, 2006). Although the sequence identity between subtypes is high, there seems to be little redundancy in their function, with evidence of subtype-specific intracellular signalling during the anti-viral response (Genin *et al.*, 2009). In mouse, IFN-α4 is the only IFN-α directly induced in response to infection with a ssRNA virus and is rapidly produced (Marie *et al.*, 1998). Similar to IFN-α4, IFN-β is also an early response gene (Marie *et al.*, 1998).

1.5.5.3. IFN-α in disease

IFN-α has been implicated in the pathogenesis of an increasing number of autoimmune diseases, including AGS, SLE, psoriasis, Sjögren's syndrome, systemic sclerosis and polymyositis (reviewed by Meyer, 2009). Therapeutic uses of IFN-α in the treatment of cancer and hepatitis can cause vascular lesions, similar to the

chilblain-like lesions seen in AGS and SLE (Bachmeyer *et al.*, 1996; Campo-Voegeli *et al.*, 1998). Additional SLE-like symptoms are also seen, most notably the development of ANA, anti-dsDNA antibodies and in rare cases humans undergoing IFN-α therapy develop active SLE (Ioannou and Isenberg, 2000). An incident of trisomy of the short arm of human chromosome 9 (the location of the type I IFN locus) was shown to result in high levels of type I IFN and the development of an SLE-like disease (Zhuang *et al.*, 2006), which further implicates IFN production in the development of an autoimmune disease state.

Mice with astrocyte-targeted overexpression of IFN-α in the CNS were shown to exhibit a significantly increased resistance to infection by neurotrophic viruses. However, they developed a structural and functional neuropathology reminiscent of that seen in AGS and congenital viral infections (Akwa *et al.*, 1998; Campbell *et al.*, 1999). This strongly implies that intracellular calcificatations and atrophy seen in AGS patients (Figure 1.4) are a direct onsequence of IFN-α production, and further suggests that activation of the innate immune system occurs in AGS patients.

1.6. How do mutations in nucleases cause AGS?

Five genes have been identified as being mutated in AGS patients, which encode two nucleases, RNase H2 and TREX1, and the putative phosphohydrolase SAMHD1 (Crow *et al.*, 2006a; Crow *et al.*, 2006b; Rice *et al.*, 2009). All identified AGS mutations in the *RNASEH2* and *TREX1* genes result in a loss of enzymatic activity, and are therefore likely to lead to an accumulation of their unprocessed endogenous nucleic acid substrates in patient cells (Figure 1.12). Self-nucleic acids are known to activate the innate immune response, inducing production of the type I IFNs, in particular, IFN-α which is known to potentiate autoimmune diseases (Chung and Older, 1997; Ioannou and Isenberg, 2000; Kataoka *et al.*, 2002; Ladoyanni and Nambi, 2005). Excessively high levels of intrathecal IFN-α is likely to be the direct cause of the white matter destruction and calcifications in the brains of individuals with AGS, and also contributes to the development of systemic features (as discussed in Section 1.5.5.3).

Evidence from the study of mice deficient in Trex1 activity suggested that Trex1 may function in the suppression of retroviral infection by degrading replication intermediates (Stetson *et al.*, 2008). Another study proposed that Trex1 acts to degrade aberrant nucleic acids accumulating in cells as replication intermediates during the cell cycle (Yang *et al.*, 2007). It is conceivable that RNase H2 could also act to degrade RNA:DNA hybrid substrates arising from both of these sources.

Although the true cellular functions of all of the AGS proteins have not yet been fully elucidated, it is clear that loss-of-function mutations in nucleases are highly likely to result in an accumulation of unprocessed nucleic acid substrates. TREX1 substrates, potentially derived from reverse transcribed endogenous retroelements, stimulate type I IFN production via recognition by an undetermined cytosolic receptor (Stetson *et al.*, 2008). The origin and nature of RNase H2 substrates that are expected to accumulate in AGS patients is unknown. Furthermore, the immunostimulatory ability of such RNase H2 substrates have not been characterised.

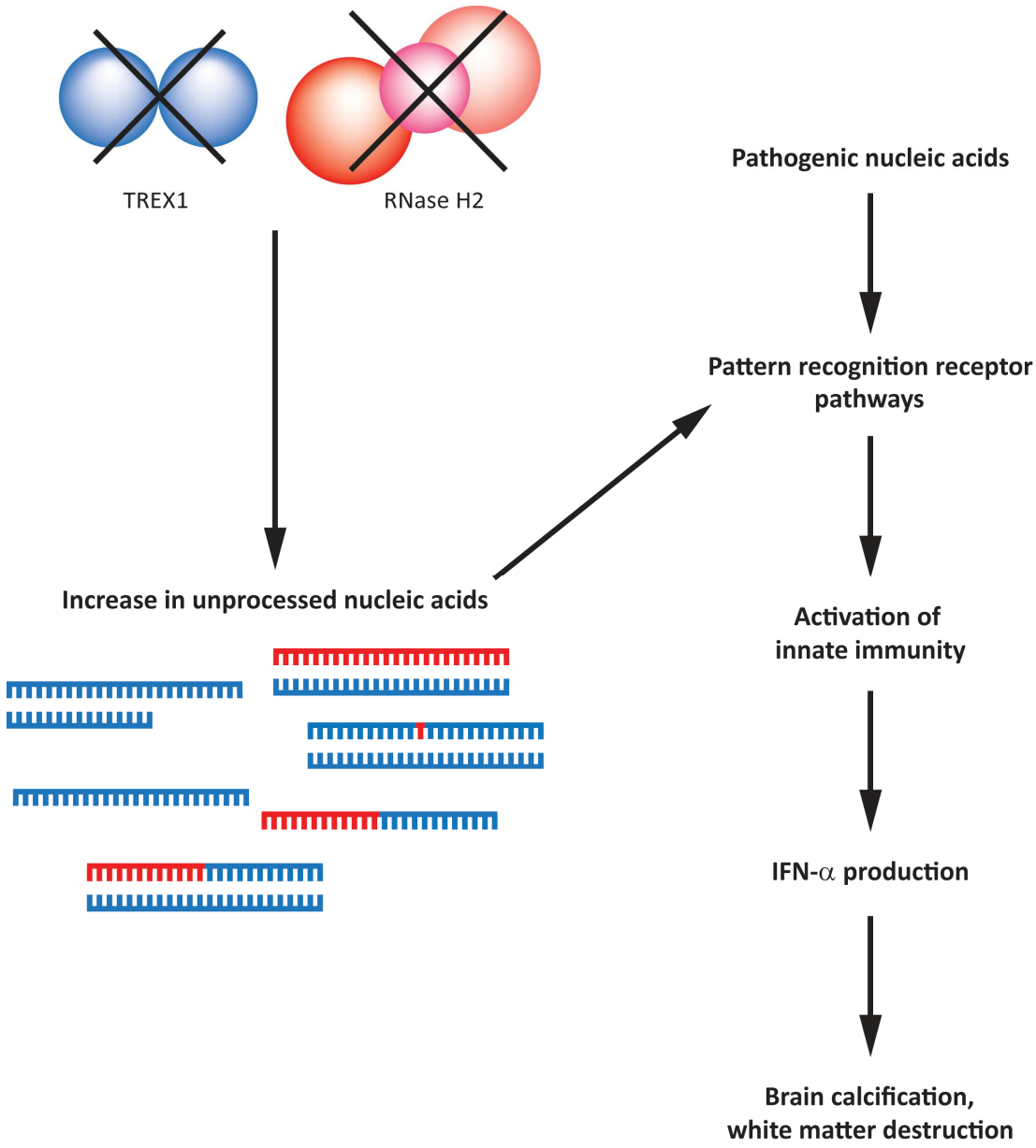


Figure 1.12. Model of how loss-of-function mutations in the AGS nucleases may result in a neuropathology reminiscent of congenital viral infection.
Impaired enzymatic activity of either TREX1 or RNase H2 is likely to result in accumulation of their unprocessed self-nucleic acid substrates. Potential endogenous substrates species are illustrated here (ribonucleotides are shown in red, deoxyribonucleotides in blue). Pathogenic nucleic acids entering a cell are detected by intracellular PRRs, triggering downstream signalling cascades and the production of IFN-α. Accummulated self-nucleic acids could potentially be inappropriately recognised by these PRRs, resulting in the development of an autoimmune response. High levels of IFN-α is likely to be a direct cause of the cerebral calcifications and atrophy in the brains of individuals with AGS and with congenital viral infection of the brain. Figure modified from Rigby *et al*. 2008.

Alternatively, the AGS proteins may function directly in the innate immune response to nucleic acids by functioning as PRRs, although this is not thought to be the case for TREX1 (Stetson *et al.*, 2008). Furthermore, they may act as components of PRR signalling pathways, required to process immunogenic self nucleic acids before the induction of type I IFN production and the subsequent damaging autoinflammatory response. Evidence of this from mice deficient in Trex1 activity led to the proposal of TREX1 as a negative regulator of cell-intrinsic immunity (Stetson *et al.*, 2008).

1.6.1. Hypothesis

Mutations which reduce/abrogate RNase H2 or TREX1 activity result in an accumulation of their unprocessed self-nucleic acid substrates which are inappropriately recognised as pathogenic by receptors of the innate immune system, and that the ensuing production of IFN-α is a direct cause of the neuropathology seen in AGS patients (Figure 1.12).

1.7. Aims of this thesis

The main aims of this thesis were to define a role for RNase H2 and its RNA:DNA hybrid substrates in innate immunity; specifically, does RNase H2 function in the innate immune response to nucleic acids and can its RNA:DNA hybrids stimulate immunity *in vitro*?

Initial experiments were performed to establish if RNase H2 is directly involved in the innate response to nucleic acids, which revealed that this was unlikely (Chapter 3). Therefore, a role for RNA:DNA hybrid substrates of RNase H2 in the activation of innate immunity was investigated. Reagents for the study of RNA:DNA hybrids were characterised (Chapter 4) before a range of synthetic RNA:DNA hybrids with different physiochemical properties were designed, synthesised, purified and validated (Chapter 5). The ability of RNA:DNA hybrids to stimulate innate immunity *in vitro* was investigated using cell lines and primary cells of the immune

system (Chapter 6). In summary, this thesis aimed to define a role for the RNase H2 complex and its nucleic acid substrates in the innate immune response and in the production of IFN-α in AGS. A greater understanding of this would provide important clues as to the cellular function of the mammalian RNase H2 complex, the development of disease pathogenesis in AGS patients, and the ability of a novel nucleic acid species to stimulate the innate immune system.

Chapter 2. Materials and Methods

2.1. General reagents

2.1.1. Sources of reagents

Unless stated otherwise, chemicals used during this thesis were obtained from Sigma-Aldrich or BDH. UltraPure^{TM} nuclease-free distilled water was from Invitrogen. Cell culture media was purchased from Gibco (Invitrogen) or Sigma-Aldrich. Oligonucleotides were synthesised by Sigma-Genosys, Eurogentec, Invitrogen and IBA. Murine cytokine qPCR primer sets were ordered from Real Time Primers Llc. All oligonucleotide sequences are listed in Appendix 1.

2.1.2. Preparing solutions

Solutions were sterilised by autoclaving at 121^{o}C for 15 min at 1 bar. Solutions unable to be autoclaved were filter sterilised through a 0.22 μm filter (Millipore). RNase-free solutions were prepared using UltraPure water, or alternatively by overnight treatment with 0.1% (v/v) diethyl pyrocarbonate (DEPC). DEPC-treated solutions were autoclaved to inactivate the DEPC. Commonly used solutions are listed in Table 2.1.

Table 2.1. Composition of commonly used buffers in this thesis

Buffer	Components
10 X PBS	1.37 M NaCl, 26.8 mM KCl, 26.5 mM $Na_2HPO_4.7H_2O$, 17.6 mM KH_2PO_4, pH 7.4
1 X TE	10 mM Tris-HCl, 1 mM EDTA, pH 8.0
20 X TBE	1.75 M Tris base, 1.78 M boric acid, 40 mM EDTA, pH 8.0
10 X TBST	0.5 M Tris-HCl pH 7.5, 1.5 M NaCl, 0.2% (v/v) Tween-20
10 X Tris-glycine SDS PAGE buffer	250 mM Tris base, 1.92 M glycine, 1% (w/v) SDS
1 X Western transfer buffer	25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS, 20% (v/v) methanol
1X Western stripping buffer	0.2 M glycine, 1% (w/v) SDS, pH to 2.5 with HCl

2.2. Cell culture

2.2.1. Human cell lines

The HeLa, HEK293 and K562 cell lines were maintained in DMEM. The CCRF-CEM cell line was maintained in RPMI-1640 containing 2 mM $_L$-glutamine, 0.2 IU/ml insulin, 1 mM oxaloacetic acid and 0.45 mM pyruvic acid. Jurkat cells were maintained in RPMI-1640 medium containing 2 mM L-glutamine and 12.5 mM MOPS. LA155n cells were maintained in a 1:1 (v/v) mixture of EMEM and Ham's F-12 containing 2 mM $_L$-glutamine. THP-1 cells were maintained in RPMI-1640 medium containing 2 mM $_L$-glutamine, 1 mM sodium pyruvate, 10 mM HEPES (pH 7.2) and 0.05 mM β-mercaptoethanol. All media were supplemented with 10% (v/v) foetal calf serum (FCS) (HyClone, Thermo Fisher Scientific), 100 U/ml penicillin and 100 μg/ml streptomycin. Trypsinisation of adherent cells was performed with trypsin:versene (1:1, v/v) at room temperature for 5 min (room temperature was between 18 – 20^oC). B lymphoblastoid cell lines (LCLs) were maintained at 2 – 5 x 10^5 cells/ml in RPMI-1640 supplemented with 15% (v/v) FCS, 100 U/ml penicillin and 100 μg/ml streptomycin. Cells were grown at 37^oC under 5% CO_2.

2.2.1.1. Differentiation of THP-1 cells into macrophage-like cells

THP-1 monocytic cells were treated with 5 ng/ml phorbol 12-myristate 13-acetate (PMA) for 48 h to induce differentiation into fully adherent macrophage-like cells, as described previously (Park *et al.*, 2007). Cells were harvested by incubation with EDTA solution (3 mM glucose, 10 mM EDTA in RPMI) for 15 min at 37°C. Differentiation into macrophage-like cells was confirmed by the upregulation of CD14 mRNA using quantitative real time-PCR (qRT-PCR).

2.2.2. Murine cell lines

The L929 cell line was maintained in DMEM containing 10% (v/v) FCS, 100 U/ml penicillin and 100 µg/ml streptomycin. Trypsinisation was performed with trypsin:versene (1:1, v/v) at room temperature for 5 min. The S9.6-secreting hybridoma cell line HB-8730 (ATCC-LGC Promochem) was maintained in DMEM supplemented with 20% (v/v) FCS, 100 U/ml penicillin, 100 µg/ml streptomycin, 0.1 mM hypoxanthine and 0.016 mM thymidine and grown at 37°C under 5% CO_2. High-density cultivation of the HB-8730 cell line was achieved using the miniPERM Classic Bioreactor System (Greiner Bio-One), following the manufacturer's instructions.

2.2.2.1. Preparation of L929-conditioned medium

4.7 x 10^5 L929 cells were seeded in a 75 cm^2 flask in 55 ml DMEM containing 10% (v/v) FCS, 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were grown under standard culture conditions for 7 days without subculturing, after which the supernatant was harvested, filtered through a 0.22 µm-filter and stored in 50 ml aliquots at –20°C.

2.2.3. Preparation and culture of primary cells

Murine primary cells were generated from wild type C57/Bl6 mice, except Rnaseh2b transgenic mice which were on a mixed background (C57/Bl6 and 129).

2.2.3.1. Mouse embryonic fibroblasts (MEFs)

Embryos were harvested at 13.5 d.p.c. (E13.5) and held in a tissue culture petri dish containing culture medium (DMEM supplemented with 20% (v/v) FCS, 100 U/ml penicillin and 100 μg/ml streptomycin). Placental and maternal tissues were removed by blunt dissection and each embryo transferred to a fresh dish of medium. Tail-tips were taken for genotyping (described in Section 2.3.2.5) and the heads and abdominal cavities were removed and discarded. The embryo bodies were separated into individual wells of a 6-well tissue culture plate (Greiner Bio-One) with 2 ml culture medium and finely chopped using a sterile scalpel. The tissues were then passed several times through a graduated pastette to further disrupt them before transfer into a 25 ml tissue culture flask with 5 ml culture medium. Adherent fibroblast colonies were visible following culture for 48 h at 37^{o}C under 5% CO_2. When cultures reached 80% confluence, they were washed with PBS and incubated with 1 ml Accutase (Innovative Cell Technologies Inc.) for 5 – 10 min at 37^{o}C to detach the cells, which were collected and pelleted by centrifugation at 300 x g for 5 min. Cells were passaged at a 1:3 ratio. MEFs were discarded after approximately four passages (prior to the cultures becoming senescent).

2.2.3.2.1. Harvesting bone marrow

Adult mice were sacrificed by cervical dislocation, and femurs and tibiae were dissected. The bones were washed once in 70% (v/v) ethanol and rinsed three times in PBS. Bone marrow was collected by flushing the marrow through with culture medium (RPMI containing 2 mM $_{L}$-glutamine, 10% (v/v) FCS, 100 U/ml penicillin and 100 μg/ml streptomycin) using a 5 ml syringe with a 25 G needle (BD Biosciences). Clumps within the bone marrow suspension were dispersed by passing the cells through a 21G needle (BD Biosciences). Cells were centrifuged at 300 x g for 5 min and resuspended to a volume appropriate for counting in culture medium.

2.2.3.2.2. Bone marrow derived-macrophages

Bone marrow was harvested as described in Section 2.2.3.2.1. Live cells were counted using trypan blue exclusion, discounting any red blood cells. They were then seeded in sterile 9 cm triple vent petri dishes (non-coated, Greiner Bio-One) at 2 x 10^5 cells/ml, in 20 ml culture medium supplemented with either 20 ng/ml M-CSF (R&D Systems) or 10% (v/v) L929 conditioned medium. Cells were cultured at 37^{o}C under 5% CO_2 for 7 days, replacing 10 ml of medium with fresh culture medium on days 4 and 6. On day 7, the medium was removed and the adherent cells washed with PBS before incubation in EDTA solution (3 mM EDTA, 10 mM glucose in RPMI) for 15 min at 37^{o}C. The suspension was collected from the plates and centrifuged at 300 x g for 15 min. The pelleted cells were resuspended in culture medium without antibiotics and viable cells counted using trypan blue exclusion. 1 x 10^6 cells were seeded in 2 ml culture medium without antibiotics into 6-well plates and the cells allowed to adhere for 4 h at 37^{o}C. Cells were washed once with medium and 0.8 ml culture medium without antibiotics was replaced in each well. Cells were stimulated with PAMPs and nucleic acid ligands (Section 2.2.4) and harvested by incubation with EDTA solution (3 mM glucose, 10 mM EDTA in RPMI) for 15 min at 37^{o}C. Analysis of expression of the surface markers CD11b and F4/80 was used to confirm the phenotype of the cells (Section 2.2.5.1).

2.2.3.2.3. Granulocyte macrophage-colony stimulating factor differentiated dendritic cells (GM-CSF DCs)

Bone marrow was harvested (Section 2.2.3.2.1) and dendritic cells (DCs) were generated as described previously (Lutz *et al.*, 1999). Cells were seeded in sterile 9 cm triple vent petri dishes (non-coated, Greiner Bio-One) at 2 x 10^6 cells per plate in 10 ml RPMI-1640 supplemented with 10% (v/v) FCS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM $_L$-glutamine and 20 ng/ml recombinant murine granulocyte macrophage colony stimulating factor (GM-CSF) (Peprotech), and were cultured at 37^{o}C under 5% CO_2. On day 3, a further 10 ml culture medium containing 20 ng/ml GM-CSF was added. On days 6 and 8, 9 ml of culture supernatant was removed and replaced with 10 ml fresh culture medium containing 20 ng/ml GM-CSF. On day 10, the cells were harvested by gentle rinsing of the plates with the

culture medium. Viable cells were counted and seeded into non-tissue culture grade sterile 12-well plates (Corning Life Sciences) at 2 x 10^6 cells per well in 0.8 ml culture medium (without antibiotics) containing 5 ng/ml GM-CSF. Cells were stimulated with PAMPs and nucleic acid ligands (Section 2.2.4) and harvested by gently rinsing each well with the culture medium. Analysis of the surface marker CD11c was used to confirm the phenotype of the cells (Section 2.2.5.1).

2.2.3.2.4. Fms-like Tyrosine Kinase 3 Ligand differentiated dendritic cells (FL-DCs)

Bone marrow cells were harvested (Section 2.2.3.2.1) and resuspended in 0.168 M NH_4Cl pH 7.2 for 1 min on ice to lyse red blood cells. Following three washes with 50 ml RPMI containing 10% (v/v) FCS, cells were resuspended in culture medium (RPMI-1640 supplemented with 10% (v/v) FCS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM $_L$-glutamine, 50 µM β-mercaptoethanol and 200 ng/ml recombinant murine Fms-like tyrosine kinase 3 ligand (Flt3-L) (Peprotech). Viable cells were counted and 2 x 10^6 cells/ml seeded in culture medium in 25 cm^3 tissue culture flasks. Following culture for 8 days at 37^{o}C under 5% CO_2, cells were harvested by gently rinsing the flask with the culture medium. Cells were centrifuged at 300 x g for 5 min and resuspended in culture medium without antibiotics. Viable cells were counted and seeded at 2 x 10^6 cells per well in non-tissue culture grade sterile 24-well plates (Corning Life Sciences) in 0.8 ml culture medium without antibiotics containing 200 ng/ml Flt3-L. Cells were stimulated with PAMPs and nucleic acid ligands (Section 2.2.4) and harvested by gently rinsing each well with the culture medium. Analysis of the surface markers CD11c, B220, CD11b and CD24 was used to confirm the phenotypes of cells (Section 2.2.5.1).

2.2.4. Stimulation of cells with PAMPs and nucleic acid ligands

Cells were stimulated with a range of TLR ligands, purchased from Invivogen unless stated. The TLR3 ligand polyinosinic:polycytidylic acid (poly (I:C)) was used at concentrations between 1 – 25 µg/ml and the TLR4 ligand lipopolysaccharide (LPS)

(Sigma-Genosys) at concentrations between 100 – 200 ng/ml. The imidazoquinoline compound R848, a TLR7/8 ligand, was used at concentrations between 1 – 2 μg/ml. The TLR9 stimulatory CpG oligonucleotide (ODN) CpG-ODN 1585, a type A CpG-ODN (5′ - GGggtcaacgttgaGGGGGG - 3′, where phosphothiorate (PS) bases are in capital letters) was used at concentrations between 1 – 10 μg/ml. 45 bp interferon stimulatory DNA (ISD) was prepared by annealing two PAGE-purified complementary phosphodiester DNA oligonucleotides (forward: 5′ - TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA - 3′, reverse: 5′ - TGTAGATCATGTACAGATCAGTCATAGATCACTAGTAGATCTGTA - 3′ (Stetson and Medzhitov, 2006). Double-stranded ISD was purified from unannealed ssDNA using size exclusion FPLC (Section 2.3.1.8). An immunostimulatory *in vitro* transcribed (IVT) RNA corresponding to the first 99 nt of the neomycin resistance gene (a gift from Delphine Goubau) (Rehwinkel *et al.*, 2010) was used to transfect IPS-$1^{-/-}$ FL-DCs at 1 μg/ml. The synthetic RNA:DNA hybrid poly (rA:dT) (Midland Certified Reagent Company, Inc.) was used at concentrations between 1 - 25 μg/ml. Nucleic acids (RNA:DNA hybrids and single stranded oligonucleotides) were used at the concentrations stated in the figure legends. All stimulations were performed in a final volume of 1 ml.

Prior to transfection, adherent cells (HeLa, L929, MEFs and macrophages) were seeded into 6-well plates in 2 ml culture medium without antibiotics and allowed to adhere for approximately 4 h at 37^{o}C. HeLa and L929 cells were seeded at 2 x 10^5 cells/ml, MEFs were seeded at 1.5 x 10^5 cells/ml and macrophages were seeded as described in Section 2.2.3.2.2. Cells were approximately 70% confluent before transfection. The medium was replaced with 0.8 ml fresh medium before the cells were transfected. Suspension cells (GM-CSF DCs and FL-DCs) were resuspended in 0.8 ml culture medium and seeded immediately before transfection (Sections 2.2.3.2.3 and 2.2.3.2.4).

For the transfection of nucleic acids into cells, Lipofectamine cationic lipid transfection reagent (Invitrogen) was used as initial pilot studies showed a high

transfection efficiency could be achieved using this reagent (Figure 6.1). Lipofectamine has been successfully used to transfect nucleic acids into a variety of cell types (see discussion in Section 6.1), therefore for continuity the same reagent was used for all transfections throughout this thesis.

Transfections were performed following the manufacturer's instructions. Two transfection solutions were used: the first contained the appropriate volume of the nucleic acid diluted in Optimem (Gibco) to a final volume of 100 µl. The second contained 2 µl Lipofectamine 2000 or Lipofectamine LTX and 98 µl Optimem. The transfection mixtures were incubated at room temperature for 5 min before they were gently added together and incubated for a further 20 min at room temperature. The mixtures were then added to 0.8 ml cells which were gently rocked to disperse the solution. Transfection control wells containing medium only, Optimem only and Lipofectamine only were included in each experiment. Cells were incubated at 37^{o}C for the time indicated in the figure legends. Subsequently, the medium from each well was harvested, centrifuged at 1500 x g for 5 min and supernatants stored at -80^{o}C until the quantity of secreted cytokines was determined by ELISA (Section 2.5.5). The cells were prepared for RNA extraction (Section 2.3.3.1), protein extraction (Section 2.4.1) or the analysis of activation marker expression by flow cytometry (Section 2.2.5.1).

2.2.5. Flow cytometry

2.2.5.1. Detection of phenotype and activation markers

Cells were harvested as described in Section 2.2, resuspended in FACS buffer (1% (v/v) FCS in PBS) and 200 µl of each sample transferred to a 96-well V-bottomed microplate (Greiner Bio-One). The plate was centrifuged at 450 x g for 2 min and the supernatants discarded. The plate was briefly vortexed to disperse the pellet before the cells were resuspended in 50 µl FcR block (anti-CD32/CD16) (final concentration 100 ng/ml; produced in-house by Andrew MacDonald's laboratory at The University of Edinburgh) diluted in FACS buffer, to block non-specific binding to Fc receptors. Cells were incubated for 10 min at 4^{o}C before 50 µl of the antibody

staining cocktail was added. All antibody dilutions used during this thesis are listed in Appendix 2. Cells were incubated for 15 min at 4^{o}C, before the plate was centrifuged as described above. Cells were washed twice in 200 µl FACS buffer, resuspended in 200 µl FACS buffer and transferred to a 5 ml polystyrene tube (BD Biosciences). Samples were acquired on a BD Biosciences FACSAriaII (macrophages and GM-CSF DCs) or a BD Biosciences FACSCantoII (FL-DCs). Data was analysed using FlowJo software (v7.6.1, Tree Star).

2.2.6. Visualisation of RNA:DNA hybrids using fluorescence microscopy

1 x 10^5 L929 cells were seeded onto a 16 mm circular glass coverslip in a 12-well tissue culture plate (Greiner Bio-Science) in 2 ml culture medium without antibiotics and left for 4 h at 37^{o}C to adhere. Cells were transfected with fluorescein-labelled RNA:DNA hybrids as described (Section 2.2.4) using Lipofectamine or other transfection reagents according to the manufacturer's instructions. The transfection solutions were applied to the cells for 4 h at 37^{o}C in the dark, after which the cells were washed once with 1.0 ml PBS and fixed using 0.5 ml 4% (w/v) paraformaldehyde (PFA) in PBS for 10 min at room temperature in the dark. Residual PFA was washed off the cells with PBS. The cell nuclei were stained with 0.05 µg of the dsDNA-specific dye 4',6-diamidino-2-phenylindole (DAPI) diluted in 0.5 ml PBS for 10 min at room temperature in the dark. The cells were rinsed in PBS, and the coverslips mounted onto glass microscope slides (SuperFrost; VWR International) using Vectashield (Vector Laboratories) and sealed. Fluorescence was visualised using a Zeiss Axioplan II fluorescence microscope using the wavelengths 360 nm (DAPI) and 488 nm (fluorescein). Image capture and analysis were performed using in-house scripts written for IPLab Spectrum (v 3.9.5, Scanalytics Corp.).

2.3. Nucleic acid methods

2.3.1. General methods

2.3.1.1. Spectrophotometric quantification

The concentration of nucleic acid samples was determined by measuring absorbance at 260 nm using a NanoDrop 1000 UV-Vis Spectrophotometer (Thermo Fisher Scientific). 1.2 µl of each sample was analysed. Samples with concentrations exceeding the limits for detection (3700 ng/µl for dsDNA, 3000 ng/µl for ssRNA and 2400 ng/µl for ssDNA, as recommended by the manufacturer) were diluted before analysis. The purity of the nucleic acid sample was determined by calculating the absorption at both 260 and 280 nm. A_{260}/A_{280} ratios for nucleic acid samples free from contaminants should be $\geq$ 1.8 (DNA) and $\geq$ 2.0 (RNA).

2.3.1.2. Calculation of the molecular weight of nucleic acids

The molecular weights (MW) of oligonucleotides were calculated using the formulas:

MW ssRNA = $(A_n \times 329.2) + (U_n \times 306.2) + (C_n \times 305.2) + (G_n \times 345.2)$
MW ssDNA = $(A_n \times 313.2) + (T_n \times 304.2) + (C_n \times 289.2) + (G_n \times 329.2)$

2.3.1.3. Ethanol precipitation

Nucleic acids were precipitated from aqueous solution by adding 2.5 volumes of ice cold absolute ethanol (200 proof, Molecular Biology Grade; Sigma) and 300 mM nuclease-free sodium acetate (NaOAc) (pH 5.2). If necessary, 20 µg nuclease-free glycogen (Molecular Biology Grade; Roche) was included as a carrier molecule. After thorough mixing by pipetting, the samples were placed at -20˚C for a minimum of 2 h. Precipitates were pelleted by centrifugation (16,000 x g, 30 min, 4˚C) and the pellet washed once with 2 volumes of 70% (v/v) nuclease-free ethanol before it was air-dried for approximately 5 - 10 min and dissolved in an appropriate volume of nuclease-free water or nuclease-free TE.

2.3.1.4. Annealing oligonucleotides

Oligonucleotides used for making RNA:DNA hybrid duplexes were subjected to HPLC or PAGE purification by the manufacturers. The purity of selected oligonucleotides was checked by denaturing PAGE analysis (Section 2.3.1.6) prior to annealing. 20 μM oligonucleotides were annealed in nuclease-free 60 mM KCl, 50 mM Tris (pH 8.0) before being denatured at 95^{o}C for 5 min. The nucleic acid mixture was then allowed to gradually cool to room temperature to permit annealing of the duplex before being aliquoted and stored at –80^{o}C or –20^{o}C. The sequences of all oligonucleotides used for making RNA:DNA hybrids are listed in Appendix 3.

2.3.1.5. Agarose gel electrophoresis

Nucleic acid samples were mixed with 1 X sample loading buffer (30% (v/v) glycerol, 0.4% (w/v) Orange G) and loaded onto horizontal 0.8% - 3% (w/v) TBE/agarose (Hi-Pure Low EEO Agarose, Biogene) gels containing 0.5 μg/ml ethidium bromide. Samples were electrophoresed in 1 X TBE at 5 V/cm. DNA markers consisting of known fragment sizes (5 bp ladder (Fermentas), 100 bp ladder (Promega) or 1 kb ladder (Invitrogen)) were included on each gel to allow estimation of fragment size. The nucleic acids were visualised using a UV transilluminator (BioDoc-It System, UVP).

2.3.1.6. Native polyacrylamide gel electrophoresis

Size separation of nucleic acids under native conditions was performed using 1.5 mm thick native polyacrylamide gels (Appendix 4). Gels were pre-run for 1 h at 5 W in 4^{o}C prior to sample loading. Samples were mixed with an equal volume of native loading buffer (30% (v/v) glycerol, 80 mM HEPES-KOH pH 7.9, 100 mM KCl, 2 mM Mg acetate, 0.03% (w/v) bromophenol blue (BPB) and 0.03% (w/v) xylene cyanol FF (XC)) and electrophoresed in 0.5 X TBE for the appropriate length of time (usually 2 – 6 h dependent on the size and strandedness of nucleic acid) at 5 W in 4^{o}C in a vertical gel tank (Cambridge Electrophoresis, Model No. EV200). BPB and XC were used as references to estimate migration distance (Appendix 4).

2.3.1.7. Denaturing polyacrylamide gel electrophoresis

Samples of nucleic acids were mixed with an equal volume of 2 x formamide loading dye (96% (v/v) deionised formamide (Sigma), 20 mM EDTA pH 8.0, 0.03% (w/v) BPB and 0.03% (w/v) XC), denatured for 5 min at 95^{o}C and immediately placed on ice to prevent renaturation. Samples were loaded onto a pre-run (1 h at 24 W) 1.5 mm thick denaturing polyacrylamide-urea gel (Appendix 4) in a vertical gel tank (Model EV200, Cambridge Electrophoresis). 2 µl of RiboRuler^{TM} Low Range RNA ladder (Fermentas) was included as a size reference. The gel was run for the appropriate length of time (usually 1 - 2 h, dependent on the size of the nucleic acid) at 24 W, with a metal plate used to distribute heat evenly. BPB and XC were used as a reference (Appendix 4). Nucleic acids in the gels were stained (Section 2.3.1.7) and imaged using a UV transilluminator (BioDoc-It System, UVP) and a phosphorimager (Model FLA-2000, Fujifilm).

2.3.1.8. Nucleic acid stains

Nucleic acids electrophoresed on polyacrylamide gels were stained with either acridine orange (BDH), ethidium bromide (BDH), Vistra Green (Invitrogen) or Sybr Gold (Invitrogen). All stains were diluted in gel running buffer (TBE) and staining was performed in a glass container in the dark at room temperature on a see-saw rocker (Stuart SSL4, Bibby Scientific) . Acridine orange staining was performed using 4.3 µM acridine orange (Lauretti *et al.*, 2003) for 10 min followed by destaining in ddH_2O for a minimum of 10 min. 0.5 µg/ml ethidium bromide was used to stain gels for 15 min, followed by destaining in ddH_2O for 15 min. Vistra Green and Sybr Gold were used according to the manufacturer's instructions (1:10,000 dilution, staining for 30 min) and the gels transferred to ddH_2O before imaging. Stained gels were imaged immediately using a UV transilluminator (BioDoc-It System, UVP) and a phosphorimager (Model FLA-2000, Fujifilm).

2.3.1.9. Fast performance liquid chromatography (FPLC) of nucleic acids

An ÄKTA FPLC™ machine (GE Healthcare) with a 24 ml Superdex 10/300 GL column (GE Healthcare) was used to perform size exclusion chromatography of nucleic acids. All FPLC was performed at 4°C. Solutions were prepared with UltraPure water and before use were passed through a 0.22 μm filter and degassed using a vacuum filter unit (MF75, Nunc). Solutions were chilled to 4°C before use. The equilibriation buffer used for each run was the same as the sample buffer the nucleic acid was contained in. Typically nuclease-free TE or oligonucleotide annealing buffer (60 mM KCl, 50 mM Tris pH 8.0) were used. Prior to beginning chromatography, the pumps of the machine were washed with 20% (v/v) ethanol before being thoroughly rinsed first with nuclease-free water and then with equilibriation buffer. The column was washed with water before being equilibriated with two column volumes (CVs) of equilibriation buffer. 100 μl of sample was loaded into into a 1 ml superloop which had been washed with equilibriation buffer prior to injection of the sample using a glass 1 ml injection needle syringe (Hamilton). The superloop was emptied with 100 μl of equilibriation buffer to inject the sample onto the column. The sample was eluted with 2 CV of equilibriation buffer at a flow rate of 0.4 ml/min. 100 μl fractions were collected into sterile, nuclease-free 96 well 2 ml boxes (MASTERBLOCK®, Greiner Bio-One). UNICORN software (GE Healthcare) was used to create run programmes and analyse data. The ÄKTA FPLC™ machine used during this thesis did not have a 230 nm or a 260 nm source fitted, therefore absorbance at 280 nm was recorded. After PAGE analysis of individual fractions, the required fractions were quantified and concentrated by ethanol precipitation (Section 2.1.3.2).

2.3.2. DNA methods

2.3.2.1. Oligonucleotide primer design

Oligonucleotides for PCR were designed using Primer3Plus (Untergasser *et al.*, 2007). Oligonucleotides for quantitative real-time PCR (qRT-PCR) were designed using Primer3Plus and the Universal ProbeLibrary Assay Design Center (Roche

Applied Science). Criteria for designing oligonucleotides included a GC content of between 40 – 60%, a melting temperature (T_m) of at least 50^{o}C and the absence of internal secondary structure. Oligonucleotides pairs were designed to have similar T_m values and were not complementary to each other at the 3′-termini, to avoid the formation of primer-dimers. Specificity of primers for qRT-PCR was validated using BLAST (Altschul *et al.*, 1990). The sequences of all oligonucleotide primers used for PCR and qPCR during this thesis are listed in Appendix 1.

2.3.2.2. Amplification of DNA by polymerase chain reaction (PCR)

Selected regions of DNA were amplified using the polymerase chain reaction (PCR). Specifically designed oligonucleotide primers (Section 2.3.2.1) were annealed to denatured template DNA and extended by a thermostable DNA polymerase. Reactions were performed in 20 µl reactions unless otherwise specified. A typical reaction mixture contained 1 X FastStart Taq DNA Polymerase reaction buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix (Invitrogen), 0.2 – 0.5 µM of each forward and reverse primer, 0.4 U FastStart Taq DNA Polymerase (Roche) and 0.1 – 100 ng template DNA. The total volume was made up to 20 µl with nuclease-free H_2O. Minus template (-DNA) and minus PCR mix (-PCR) controls were included with each experiment. A standard PCR programme was used to achieve amplification, which consisted of an initial denaturing step at 95^{o}C for 6 min, then 30 cycles of denaturing at 95^{o}C for 30 s, annealing at X^{o}C for 45 s, extension at 72^{o}C for 1 min and a final extension at 72^{o}C for 10 min, where X indicated the optimal annealing temperature of for each set of oligonucleotide primers, as determined by a temperature gradient. Thermal cycling was performed using a DNA Engine Tetrad 2 thermal cycler (MJ Research). PCR products were analysed by agarose gel electrophoresis (Section 2.3.1.4).

2.3.2.3. High fidelity PCR

When high accuracy amplification of the template DNA was required, such as the production of dsDNA templates for *in vitro* transcription (IVT), PCR was performed using Phusion Flash High-Fidelity PCR Master Mix (Finnzymes) which contains a

modified proof-reading DNA polymerase with a highly processive activity. Reactions consisted of 1 X Phusion Flash High-Fidelity PCR Master Mix, 0.5 µM of each primer and 1 – 100ng of template DNA. Thermal cycling conditions consisted of initial denaturation of the template at 98^{o}C for 10 s, then 30 cycles of denaturing at 98^{o}C for 1 s, annealing at 66^{o}C for 1 s, extension at 72^{o}C for 3 s and a final extension at 72^{o}C for 1 min. The annealing temperature and extension time were optimised for each set of oligonucleotide primers. Thermal cycling was performed using a DNA Engine Tetrad 2 thermal cycler (MJ Research). PCR products were analysed by agarose gel electrophoresis (Section 2.3.1.4).

2.3.2.4. Amplification of genomic DNA for genotyping of mice

Genomic DNA (gDNA) obtained from Trizol extraction of embryonic nucleic acids (Section 2.3.3.1.2) was amplified using a Genomiphi kit (GE Healthcare). 1 µl of DNA was incubated with 9 µl substrate buffer for 3 min at 95^{o}C before 1 µl of DNA polymerase mix and 9 µl of reaction buffer was added on ice. The reaction was incubated at 30^{o}C for 16 h before it was heated to 65^{o}C for 10 min to deactivate the enzyme. 1 µl of the amplified DNA was used as a template for the genotyping PCR (Section 2.3.2.5).

2.3.2.5. Genotyping of transgenic mice with mutations in *Rnaseh2b*

PCR was performed using 1 µl of template gDNA or boiled embryonic tail soup in a 25 µl reaction containing 12.5 µl 2 X ReddyMix PCR master mix (Thermo Scientific) and 1 µM of each oligonucleotide (loxp_forw, exon7_rev and neo_rev). PCR conditions were as follows: denaturation at 94^{o}C for 2 min followed by 30 cycles of 45 s at 94^{o}C, 45 s at 58^{o}C, 45 s at 72^{o}C and a final extension step at 72^{o}C for 5 min. PCR products were electrophoresed using a 1% (w/v) agarose gel (Section 2.3.1.4). Expected product sizes were 220 bp for wildtype, 220 bp and 420 bp for heterozygotes and 420 bp for homozygous mutants.

2.3.2.6. Restriction digest of PCR products

15 pmol of PCR product generated with a reverse primer containing an EcoRI restriction site was incubated with 3 U *Eco*RI (Roche Applied Science) for 1 h at 37^{o}C. As the digested product will migrate at a similar size to undigested product on an agarose gel, plasmid DNA containing EcoRI sites was included as a control for digestion by the enzyme. Digested DNA was purified as described in Section 2.3.2.8.

2.3.2.7. Lambda exonuclease digest of PCR products

188 bp PCR products generated using a 5′-phosphorylated forward primer were digested with λ-exonuclease (M0262, New England BioLabs) which degrades the phosphorylated strand of a DNA duplex, leaving the non-phosphorylated strand intact. The concentration of the DNA and the enzyme as well as the reaction time was optimised. The digestion of 100 ng/µl of DNA with 5 U λ-exonuclease in the reaction buffer supplied with the enzyme for 20 min at 37^{o}C was found to result in the highest yield of ssDNA product.

2.3.2.8. Purification of PCR products

To removed contaminating oligonucleotides, dNTPs, salts and proteins from PCR products and restriction enzyme-digested DNA longer than 100 bp, a QIAquick PCR Purification Kit (Qiagen) was used according to the manufacturer's instructions. DNA was eluted into 30-50 µl of nuclease-free water or buffer EB (10 mM Tris-HCl, pH 8.5) and stored at -20^{o}C. PCR products and digested DNA less that 100 bp in length and those requiring concentration were purified using a ChIP DNA Clean & Concentrator Kit (Zymo Research) and eluted in 6-30 µl of nuclease-free water or the DNA elution buffer provided by the manufacturer. Alternatively, PCR products or digests were separated by size by agarose gel electrophoresis and gel purified using a QIAquick Gel Extraction Kit according to the manufacturer's instructions and eluted with 30-50 µl of nuclease-free water or buffer EB (10 mM Tris-HCl, pH 8.5).

2.3.2.9. Techniques for generating long ssDNA

A number of techniques to generate 188 nt ssDNA were used during this thesis and are presented in Chapter 5. Specific methodology is described in this section.

2.3.2.9.1. Asymmetric PCR

A 362 bp region of the CMV promoter was amplified using 1 ng of an endotoxin-free maxiprep of pEGFP-N1 (Clontech) as described in Section 2.3.2.2, using the oligonucleotide primers CMVprom_F and CMVprom362_R. The PCR product was purified as described in Section 2.3.2.8 and used as a template for asymmetric PCR. PCR amplification was performed in 20 µl volumes using 1X FastStart Taq DNA Polymerase reaction buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix (Invitrogen), 0.5 µM forward primer (CMVprom_F), 5 µM reverse primer (CMVprom188_R), 0.4 U FastStart Taq DNA Polymerase using 0.1 – 10 ng template DNA. PCR conditions were as specified in Section 2.3.2.2, using an annealing temperature of 62^{o}C and varying the number of cycles from 10 to 50.

2.3.2.9.2. Reverse primer only PCR

PCR was performed on a 362 bp template, as described in Section 2.3.2.9.1 with some modifications. The amount of template, concentration of oligonucleotide primer, annealing temperature and number of cycles were optimised. Reactions using 60 ng of template per 20 µl reaction with 0.5 µM of reverse primer (the forward primer was omitted from the reaction), gave the best yield of product (as discussed in Chapter 5).

2.3.2.9.3. Biotinylated dsDNA experiments

A 188 bp region of the CMV promoter was amplified using a biotinylated forward oligonucleotide primer (biotin-CMVprom_F) and an unmodified normal reverse oligonucleotide primer (CMVprom188_R). PCR conditions were as specified in Section 2.3.2.2. The PCR product was gel purified as described in Section 2.3.2.8 before being used in subsequent applications.

2.3.2.9.4. Streptavidin supershift to purify the non-biotinylated strand

500 ng of biotinylated dsDNA PCR product was incubated in the presence or absence of 2 μl of Streptavidin-Horseradish Peroxidase (HRP) conjugate (GE Healthcare), in a volume of 50 μl in 1.5ml nuclease free tubes. The tubes were rotated at 20 rpm (Stuart SB3 tube rotator, Bibby Scientific) for 30 min at room temperature. Non-biotinylated PCR products of the same length were included as controls. Half of each sample was heated to 95^{o}C for 5 min before 200 ng of each sample was analysed using native PAGE (Section 2.3.1.5).

2.3.2.9.5. Binding to streptavidin beads

100 μl of resuspended streptavidin magnetic beads (Dynabeads® M-280, Invitrogen) were washed three times in the binding and washing buffer (5 mM Tris pH7.5, 0.5 mM EDTA, 1 M NaCl, nuclease-free) before they were incubated with 2 μg of biotinylated 188 bp DNA diluted in binding and washing buffer. The mixture was rotated at 10 rpm (Stuart SB3 tube rotator, Bibby Scientific) for 30 min at room temperature to allow complexing of the biotinylated DNA with the streptavidin beads. The beads were immobilised using a magnetic stand (Dynal, Invitrogen). The non-bound fraction (NBF) was removed and the beads washed three times in binding and washing buffer. To recover the non-biotinylated strand using heat denaturation, the beads were resuspended in 20 μl TE and heated to 90^{o}C for 5 min to denature the strands. The ssDNA-containing TE was removed immediately, to prevent reannealing of the single-strands. To recover the non-biotinylated strand by alkali denaturation, samples were incubated with 100 μl 0.1 M NaOH, 0.2 M NaCl for 10 min at room temperature on at 10 rpm (Stuart SB3 tube rotator, Bibby Scientific). The supernatant was recovered and the treatment repeated two further times. The three resulting supernatants were pooled. The biotinylated strand was recovered from the beads by boiling for 2 min at 90^{o}C in 40 μl 95% (v/v) deionised formamide, 10 mM EDTA to break the biotin-streptavidin bonds. All fractions were concentrated by ethanol precipitation and analysed by electrophoresis.

2.3.2.9.6. Spin columns to separate single- and double-stranded nucleic acids

The use of spin columns to separate ss- and ds- nucleic acids based upon their differential binding to silica under certain chemical conditions was performed using methods previously described by Borodina *et al*. and Beld *et al*. (Beld *et al.*, 1996; Borodina *et al.*, 2003). All centrifugation steps were performed 16,000 x g for 1 min at room temperature. A mixture of λ-exonuclease digested and undigested 188 bp DNA (total 10 μg in a volume of 100 μl) was diluted with 5 volumes of dsDNA binding buffer (5 M guanidine thiocyanate (GuSCN), 100 mM EDTA, pH 8.0). This was loaded onto a spin column with a silica membrane (Column 1) (spin column from a QIAquick PCR Purification Kit, Qiagen) and centrifuged. The ssDNA-containing flow through was mixed with 0.5 volumes of ssDNA binding buffer (5 M GuSCN, 100 mM 2-(N-Morpholino)ethanesulfonic acid (MES) pH 5.0 – 6.0, 250 mM $MgCl_2$, 0.5% (v/v) Triton X-100) and loaded onto a second silica spin column (Column 2) which was centrifuged. Column 1 (containing bound dsDNA) was washed with 400 μl dsDNA binding buffer before both spin columns were washed once with 400 μl wash buffer (10 mM Tris-HCl pH 7.5, 80% (v/v) ethanol). The membranes were dried by centrifuging once again without buffer before DNA was eluted by incubating each silica membrane in 30 μl of 10 mM Tris pH 8.0 for 1 min and collecting the flow through in a nuclease-free 1.5 ml tube. dsDNA would be expected to elute from Column 1 and ssDNA from Column 2.

2.3.2.9.7. Hydroxyapatite column chromatography

The separation of ds- and ss-DNA using hydroxyapatite was performed using a method based upon that described by Sambrook and Russell (Sambrook and Russell, 2000). Chromatography was performed at room temperature. A 1 M stock solution (pH 6.8) of sodium phosphate buffer was prepared from 1 M NaH_2PO_4 and 1 M Na_2HPO_4. A series of buffers of decreasing concentrations was then prepared using this stock solution. 1 g of Bio-Gel HTP hydroxyapatite powder (130-0420, Bio-Rad Laboratories) was rehydrated with 2 ml 50 mM phosphate buffer. The slurry was transferred to an empty 20 ml plastic chromatography column (Econo-Pac, Bio-Rad

Laboratories) and allowed to settle. The column volume (CV) was adjusted to 1 ml using 50 mM phosphate buffer before the resin was washed with 10 CVs of 50 mM phosphate buffer. A mixture of λ-exonuclease digested and undigested 188 bp DNA (total 50 μg) was diluted with 50 mM phosphate buffer to final volume of 1 ml and added to the column. The DNA solution was held in the column for 10 min to allow binding to the hydroxyapatite before it was allowed to flow through the column at a flow rate of 6 ml/h. This non-bound fraction (NBF) was retained for analysis. Fractions were collected in 15 ml nuclease-free polypropylene tubes. The column was washed with 4 CV of 50 mM phosphate buffers, the first 1 ml of which was collected (wash fraction). 1 CV each of increasing concentrations of phosphate buffers (ranging from 75 mM to 400 mM) was added to the column and each fraction collected. Fractions were dialysed into ddH_2O using dialysis tubing (SnakeSkin, Thermo Fisher Scientific), changing the ddH_2O once an hour for 3 h, before they were concentrated using ethanol precipitation (Section 2.3.1.2) using 20 μg glycogen as a carrier molecule. Pellets were resuspended in 10 μl TE and quantified using a NanoDrop before analysis by native PAGE (Section 2.3.1.5).

2.3.3. RNA methods

RNA was extracted from cells, tissues and embryos and was quantified using a NanoDrop spectrophotometer (Section 2.3.1.1). RNA quality was assessed by electrophoresis on a 1% (w/v) agarose gel containing 0.5 μg/ml ethidium bromide. All procedures using RNA were performed using filter pipette tips. RNA was stored at -80^{o}C.

2.3.3.1. Extraction of RNA from cells

RNA was isolated from cells using an RNeasy Mini Kit (Qiagen). Medium from monolayer cells was discarded and the cells lysed directly in the culture vessel using an appropriate volume of the supplied lysis buffer (RLT). Suspension cells were harvested and pelleted by centrifugation at 300 x g for 5 min before they were resuspended in an appropriate volume of buffer RLT. Cells were homogenised using

a Qiashredder column (Qiagen) by centrifugation at 16,000 x g for 2 min. The homogenate was mixed with an equal volume of 70% (v/v) ethanol and applied to a spin column. RNA was purified according to the manufacturer's protocol, including a treatment with 30 U RNase-free DNase I (Qiagen) on the column for 15 min at room temperature to degrade any DNA. All centrifugation steps were performed at 10,000 x g for 1 min at room temperature. RNA was eluted into 50 µl RNase-free water.

2.3.3.2. Extraction of RNA from tissues

Brains, spleens, thymii and hearts were dissected from adult mice, placed in 2 ml RNase-free tubes (Ambion), snap frozen in liquid nitrogen and stored at –80^{o}C until RNA was extracted. Larger tissues were roughly chopped using an RNase-free blade before freezing to allow better penetration of the TRIzol reagent (Invitrogen). Tissues were thawed on ice in 1 ml TRIzol before they were disrupted and homogenised using a rotor-stator homogeniser (Polytron PT 3100, Kinematica) using a 12 mm dispersing aggregate fitting (PT-DA 3012/2EC, Kinematica) at 25,000 – 30,000 rpm for 10 s. Between three and six repeats were required to fully homogenise the tissues, with samples placed on ice between repeats. 500 µl of homogenate was transferred to an RNase-free 1.5 ml tube and mixed with 1 volume of TRIzol. Samples were then centrifuged at 12,000 x g for 10 min at 4^{o}C to pellet any insoluble material. The RNA-containing supernatant was transferred to a new 1.5 ml tube. 200 µl chloroform was added to each sample and the tubes shaken vigorously by hand for 15 s. Following incubation at room temperature for 2 min, samples were centrifuged at 12,000 x g for 15 min at 4^{o}C to separate the phases. The RNA-containing aqueous phase was removed and transferred to a new 1.5 ml tube. To precipitate the RNA, 500 µl isopropanol was mixed with the sample. Following a 10 min incubation at room temperature, the sample was centrifuged at 12,000 x g for 10 min at 4^{o}C. The pellet was washed with 1 ml 70% (v/v) RNase-free ethanol before being air-dried for 5 – 10 min and dissolved in 100 µl nuclease-free water with a 10 min incubation at 55 ^{o}C to aid solubilisation.

2.3.3.3. Extraction of RNA from embryos

Embryos were dissected and immediately frozen on dry ice before storage at -80^{o}C until RNA was extracted. Embryos were thawed on ice in TRIzol (200 µl for E8.5, 500 µl for E9.5) and homogenised by passing ten times through a series of sterile needles, decreasing in bore size (19G, 23G, 25G) attached to a sterile 1 ml syringe. The samples were incubated at room temperature for 5 min to allow complete dissociation of nucleoprotein complexes before 0.2 volumes of chloroform was added to each sample and the tubes shaken vigorously by hand for 15 s. Following incubation for 2 min at room temperature, samples were centrifuged at 12,000 x g for 15 min at 4^{o}C. The aqueous phase was removed and transferred to a new 1.5 ml tube. To precipitate the RNA, 0.5 volumes of isopropanol and 20 µg glycogen were mixed with the sample and incubated for 10 min at room temperature. After centrifugation at 12,000 x g for 10 min at 4^{o}C, the supernatant was discarded and the pellet washed with 1 volume of 70% (v/v) RNase-free ethanol. The RNA pellets were air dried for 5 – 10 min at room temperature before being dissolved in 30 µl nuclease-free water with incubation at 55^{o}C for 10 min. To obtain DNA for genotyping, 0.3 volumes of 100% (v/v) ethanol were added to the organic/interphase solution following removal of the aqueous phase. The DNA was precipitated by incubating the mixture at room temperature for 3 min, followed by centrifugation at 2000 x g for 5 min at 4^{o}C. The supernatant was removed and the pellet resuspended in 5 µl of TE, and either used directly as a template for PCR (Section 2.3.2.5) or, if insufficient DNA was obtained, as a template for amplification of the DNA (Section 2.3.2.4).

2.3.3.4. Clean up and DNase I treatment of TRIzol-extracted RNA

The volume of RNA extracted from tissues and embryos using TRIzol was adjusted to 100 µl using RNase-free water before mixing with 350 µl buffer RLT from the RNeasy Mini Kit. 200 µl absolute ethanol was mixed with the RNA and the whole sample applied to a spin column and centrifuged at 10,000 x g for 15 s. All centrifugation steps were performed at room temperature. The column was washed once with 350 µl wash buffer RW1 before incubation with 30 U RNase-free DNase I for 15 min at room temperature. The column was then washed once with RW1, twice

with buffer RPE and then the membrane dried by centrifugation at 12,000 x g for 1 min. The RNA was eluted into 50 µl nuclease-free water.

2.3.3.5. Reverse transcription of RNA

First strand complementary DNA (cDNA) synthesis was performed in an RNase-free 0.5 ml tube using 1 µg of each RNA sample in a 14 µl reaction containing 40 U Protector RNase Inhibitor (Roche), 100 pmol random primers (Promega), 5 mM RNase-free DTT made up to volume with nuclease-free water. The reaction was incubated at 70^{o}C for 5 min to melt RNA secondary structures before being cooled on ice for 5 min to prevent secondary structures reforming. After a brief centrifugation step to collect condensation from the lids of the tubes the reverse transcription mix was added (1 µM dNTPs, 20 U AMV Reverse Transcriptase (Roche), 1 X AMV-Reverse Transcriptase buffer, made up to 6 µl with nuclease-free water). The reaction was incubated at 42^{o}C for 60 min before the enzyme was deactivated at 75^{o}C for 8 min. As a control for gDNA contamination of the cDNA, a minus reverse transcriptase (–RT) control was prepared for every cDNA sample by omitting the AMV reverse transcriptase. cDNA was stored at -20^{o}C.

2.3.3.6. Quantitative real-time PCR (qRT-PCR)

Gene expression was measured using quantitative real-time PCR (qRT-PCR) using Sybr Green which binds dsDNA PCR products, emitting a fluorescent signal upon excitation. Sybr Green does not distinguish between specific and non-specific products, therefore to help prevent contamination of the signal by primer-dimers, each oligonucleotide primer set was optimised to determine the lowest possible concentration that could be used using a titration curve. Where possible, intron-spanning primers were designed to avoid amplification of gDNA.

Reactions were performed in a 10 µl volume in clear V-bottomed 384 well PCR plates (ABgene) with optically clear plate seals (ABgene). A typical reaction consisted of 1 µl template cDNA, 1 X Brilliant II Sybr Green qPCR Master Mix (Stratagene), 0.3 µM passive reference dye (ROX) and 0.2 µM of each oligonucleotide primer in nuclease-free water. Each sample and -RT control were

analysed in triplicate and a no template control (NTC) was included for each master mix. A housekeeping gene was included as an endogenous control for each sample. *GAPDH* (encoding glyceraldehyde 3-phosphate dehydrogenase) and *Actb* (encoding β-actin) were found to be suitable endogenous controls for the experiments performed during this thesis as their expression did not alter following treatment of cells with ligands of the innate immune system.

Quantitative real-time PCR was performed using an ABI Prism HT7900 Sequence Detection System (Applied Biosciences) with the following protocol: 2 min at 50^{o}C, 10 min at 95^{o}C followed by 40 cycles of 15 s at 95^{o}C, 1 min at 60^{o}C. The 7900HT continuously detected the fluorescence of each well in the plate in real time, producing a cycle threshold (C_T) value for each well. This is defined as the cycle when the fluorescence detected is statistically significant above background. To confirm that the fluorescence detected was true amplification and not as a consequence of non-specific product formation, dissociation curve analysis was performed on the PCR products. Specificity of the oligonucleotide primers was also confirmed by electrophoresis of the PCR products on an agarose gel (Section 2.3.1.4) to ensure a single PCR product of the correct size was generated. The C_T values for each sample were used to calculate the expression of each target gene normalised to the housekeeping gene using the comparative C_T method (Livak and Schmittgen, 2001). The relative expression of target genes in treated versus untreated samples was represented as fold mRNA induction.

2.3.3.7. Microarray analysis of gene expression

Analysis of gene expression in transgenic mice with mutations in *Rnaseh2b* was performed using whole-genome gene expression arrays.

2.3.3.7.1. Amplification and biotin-labelling of RNA samples

Total RNA was extracted from cell and tissue samples as described previously in this section. The integrity of the RNA was evaluated by agarose gel electrophoresis (Section 2.3.1.4) and on RNA 6000 Nano Chips using a 2100 BioAnalyzer (Agilent

Technologies). Samples with an RNA Integrity Number (RIN) of less than 7 were discarded. The RNA was amplified and biotinylated using an Illumina® TotalPrep™ RNA Amplification Kit (Ambion), according to the manufacturer's instructions, using 200 ng (for embryos) or 300 ng (MEFs and brain) of input RNA. The *in vitro* transcription step was performed for 14 h, incubating at 37°C on a DNA thermal cycler. The concentration of the purified amplified antisense RNA (cRNA) was determined using a NanoDrop (Section 2.3.1.1), and the size distribution analysed using a 2100 BioAnalyzer. cRNA bioanalyzer profiles consisted of a distribution of sizes from 250 to 5500 nt.

2.3.3.7.2. Whole-genome expression arrays

Whole-genome gene expression analysis was performed using an Illumina Beadstation at the Wellcome Trust Clinical Research Facility, Edinburgh. 1.5 μg of each cRNA sample was loaded onto a MouseWG-6 v2.0 Expression BeadChip, which contains 45,281 probes corresponding to over 33,000 annotated coding transcripts, including retrotransposon sequences. The BeadChips contain inbuilt controls for detecting noise and background signal across the chip, and the average signal for a set of housekeeping genes and hybridisation controls. Biological triplicates for each sample were included.

2.3.3.7.3. Data analysis

Microarray data was analysed with the statistical programming language R, using its beadarray (Dunning *et al.*, 2007) and Limma (Smyth, 2004) packages by Graeme Grimes (MRC HGU, Edinburgh). Raw, non-normalised bead-summary values were imported from the Illumina BeadStudio software into R using the beadarray package. Quantile normalisation was applied to the data to enable comparison between arrays. A linear model was applied to the expression data for each gene. To determine statistically differentially expressed genes the results of the linear model were summarised and a Bayes moderated *t*-test applied. To control for multiple testing, a Benjamini and Hochberg false discovery rate *p*-value of < 0.05 was used to identify significantly differentially expressed genes.

Gene lists were imported into the DAVID (Dennis *et al.*, 2003) and FunNet (Prifti *et al.*, 2008) web-accessible programs which group genes according to their biological functions as described in genomic databases, including GenBank (Benson *et al.*, 2008) and Gene Ontology (Ashburner *et al.*, 2000). This allowed the expression profiles of each dataset to be represented in terms of the biological roles of each transcript, and biological pathways significantly altered in expression to be identified.

2.3.3.8. *In vitro* transcription (IVT)

Templates for *in vitro* transcription (IVT) were prepared by amplifying regions of the CMV promoter from pEGFP-N1 (Clontech). Each forward oligonucleotide primer containing the minimal T7 promoter sequence (5'-TAATACGACTCACTATAGG-3') followed by the first 20 nucleotides of the sequence. The 3' end of the reverse oligonucleotide primer contained the sequence for an *Eco*RI restriction site (5'-GAATTC-3'). The PCR, restriction enzyme treatment and purification of the template were performed as described in Section 2.3.2. IVT was performed using a MEGAshortscript™ High Yield Transcription Kit (Ambion) to obtain maximal yields of RNA transcript. Reaction mixtures were assembled at room temperature and contained 100-200 ng template DNA, 7.5 mM rNTPs (ATP, CTP, GTP, UTP), 2 µl 10 X T7 reaction buffer and 2µl T7 enzyme mixture, made up to a reaction volume of 20 µl with nuclease-free water. Following incubations at 37^{o}C for 2 h in a hybridisation oven, the mixture was treated with 5 µl Turbo DNase (provided by the manufacturer) for 1 h at 37^{o}C to remove all traces of the DNA template. The enzymatic reaction was terminated by diluting the reaction with 115 µl nuclease-free water and adding 15 µl Ammonium Acetate Stop Solution (5M ammonium acetate, 100 mM EDTA). The RNA was the precipitated using 2.5 volumes of ice cold 100% (v/v) ethanol (Section 2.3.1.2.) The RNA transcript was analysed by electrophoresis on a denaturing polyacrylamide-urea gel (Section 2.3.1.6.).

2.3.3.9. Purification of ssNA and RNA:DNA hybrids from denaturing polyacrylamide gels

Following electrophoresis on a denaturing polyacrylamide-urea gel (ssRNA and ssDNA) (Section 2.3.1.6) or a native polyacrylamide gel (RNA:DNA hybrids) (Section 2.3.1.5), nucleic acids were visualised by UV shadowing using a hand-held UV lamp (UVITech Ltd.) at 254 nm and a fluor-coated TLC plate (Ambion). Alternatively, gels were stained with Sybr Gold (see Section 2.3.1.7) and nucleic acids visualised using a using a UV Transilluminator (UVItec). The desired bands were excised using a sterile RNase-free blade, placed into a centrifuge tube with a 0.22 µm cellulose acetate filter (Spin-X, Costar) and crushed using an RNase-free pipette tip. The nucleic acid was eluted by passive diffusion with three incubations in 200 µl RNA elution buffer (10 mM Tris-HCl pH 7.6, 0.1% (w/v) SDS, 1 mM EDTA, pH 8.0) each for 2 h at 40 rpm (Stuart SB3 tube rotator, Bibby Scientific) at room temperature. The nucleic acid-containing buffer was recovered by centrifugation of the column at 16,000 x g for 3 min. The three buffer fractions were pooled and ethanol precipitated as described in Section 2.3.1.2, but NaOAC was replaced with 150 mM NaCl, to avoid precipitation of SDS. Ethanol precipitation removes > 97% of Sybr Gold from nucleic acids, according to the manufacturer (Invitrogen). Nucleic acids were resupended in an appropriate volume of nuclease-free water.

2.3.4. RNA:DNA hybrid methods

2.3.4.1. Reverse transcription to generate an RNA:DNA hybrid

A region of the CMV promoter sequence was amplified and used as a template for IVT, as described in Sections 2.3.2.3 and 2.3.3.8. The RNA transcript was gel purified (Section 2.3.3.9) and used as a template for reverse transcription. Reactions were performed in 0.5 ml RNase-free tubes in 35 µl reactions each containing 2.5 µg RNA, 100 U Protector RNase Inhibitor, 25 pmol reverse strand specific primer (CMVprom188EcoRI_R, which contains extra bases at the 5′-end to accommodate the overhang left on the RNA following *Eco*RI digest) made up to volume with nuclease-free water. The reaction was incubated at 70^{o}C for 5 min to melt RNA secondary structures before being cooled on ice for 5 min. After a brief

centrifugation step to collect condensation from the lids of the tubes, the reverse transcription mix was added (1 µM dNTPs, 500 U M-MLV Reverse Transcriptase (RNase H point mutant) (Promega), 1 X M-MLV-Reverse Transcriptase buffer, made up to a final volume of 15 µl with nuclease-free water). The reaction was incubated at 42^{o}C for 60 min before the enzyme was deactivated by heating at 75^{o}C for 8 min. A minus enzyme (-RT) control was included. Multiple reactions were pooled together and concentrated by ethanol precipitation. RNA:DNA hybrids were quantified using a NanoDrop (using the A_{260} conversion for dsDNA) and were verified as described in Chapter 5.

2.3.4.2. Detection of RNA:DNA hybrids using dotblots

20 pmol of nucleic acids in a volume of 1µl were dotted onto a nitrocellulose membrane (Hybond-ECL, GE Healthcare) and UV crosslinked to the membrane using a Stratalinker (Stratagene) at 1200 mJ/cm^2 for 30 s. The membrane was blocked in antibody buffer (2% (w/v) bovine serum albumin (BSA) in PBS) for 1 h at room temperature. Following three 5 min washes in PBS, the membrane was incubated with concentrated hybridoma supernatant containing the RNA:DNA hybrid-specific antibody S9.6 diluted 1:2000 in antibody buffer overnight at 4^{o}C. To detect ssNA, membranes were incubated with a monoclonal anti-ssDNA antibody (Clone F7-26, Millipore) diluted 1:500 in antibody buffer. The membrane was washed three times with PBS containing 0.1% (v/v) Tween-20 and incubated with horseradish peroxidise (HRP)-conjugated goat anti-mouse IgG antibody (Dako Corporation), diluted 1:5000 in antibody buffer containing PBS containing 0.1% (v/v) Tween-20 for 1 h at room temperature. The membrane was again washed three times with PBS containing 0.1% (v/v) Tween-20 to remove any unbound secondary antibody before visualisation of the antigen-antibody complexes using enhanced chemiluminescence (ECL) Western blotting reagents (GE Healthcare) with exposure to photographic film (BioMax Film, Kodak).

2.3.4.3. Immunodetection of electrophoresed RNA:DNA hybrids

Nucleic acids were electrophoresed on a 1 mm thick native polyacrylamide gel (Section 2.3.1.5). The gel was stained with Sybr Gold (Section 2.3.1.7) and imaged

using a UV transilluminator. The gel was rinsed in ddH_2O for 10 min before nucleic acids were transferred to a positively charged nylon membrane (Amersham Hybond-N+, GE Healthcare) using a Trans-Blot® Semi-Dry Electrophoretic Transfer Cell (Bio-Rad Laboratories). The gel and membrane were sandwiched between two sheets of thick (0.92 mm) Whatman® blotting paper soaked in 0.5 X TBE and electrophoresed for 30 min at constant voltage of 20 V. This ensured that the current did not exceed 3 mA/cm^2 and so prevented an excessive build up of heat during the transfer. Post-transfer, 20 pmol of an 18 bp RNA:DNA hybrid was dotted onto the corner of the membrane as a positive control for the antibody and the nucleic acids crosslinked to the membrane as described in Section 2.3.4.2. To confirm transfer of nucleic acids onto the membrane, it was stained with Sybr Gold (Section 2.3.1.7) and nucleic acids visualised using a hand-held UV lamp (UVITech Ltd.) at 254 nm with a fluor-coated TLC plate (Ambion). In addition, the gel was imaged using a UV transilluminator post-transfer to confirm nucleic acids had transferred to the membrane and were no longer present in the gel.

The membrane was rinsed in ddH_2O for 10 min before it was blocked in 1% (w/v) blocking agent (GE Healthcare) diluted in PBS containing 0.1% (v/v) Tween-20 overnight at 4^oC to prevent any non-specific binding. Following three washes in wash buffer (PBS containing 0.1% (v/v) Tween-20), the membrane was incubated with 400ng of protein A/G-purified S9.6 which was diluted in 2 ml of 2% (w/v) BSA/PBS overnight at 4^oC. The membrane was washed three times in wash buffer before being incubated with HRP-conjugated goat anti-mouse IgG antibody diluted 1:5000 in PBS containing 2% (w/v) BSA and 0.1% (v/v) Tween-20 for 1 h at room temperature. Unbound secondary antibody was washed away with three washes in wash buffer and the antigen-antibody complexes were visualised using enhanced chemiluminescence (ECL) Western blotting reagents (GE Healthcare) by exposure to photographic film (BioMax Film, Kodak).

2.3.4.4. Multiwell plate assay to detect RNA:DNA hybrids

18-mer nucleic acids were diluted to 5 ng/µl in nuclease-free 10 mM Tris-HCl pH 8.0 to a final volume of 44 µl. An equal volume of the DNA coating agent Reacti-

Bind (Pierce) was added to each nucleic acid. Controls using 10 mM Tris-HCl instead of React-Bind were included, as was a Reacti-Bind only control. Each sample was transferred to a glass 4 ml shell vial (Thermo Fisher Scientific) and mixed for 10 min at room temperature using a see-saw rocker (Stuart SSL4, Bibby Scientific) at 40 rpm. The nucleic acid-Reacti-Bind complexes were added to a flat-bottomed 96-well plate (MaxiSorp®, Nunc) (final volume 88 µl/well) and the plate was incubated on the rocker (10 rpm) at room temperature for 2 h, to allow bonding of the nucleic-acid to the plate. Each nucleic acid was added to the plate in triplicate. The plate was washed five times with PBS to remove any unbound nucleic acid before purified S9.6 antibody (200 ng/µl, diluted 1:1000) in 1% (w/v) BSA in PBS was added to each well (50 µl/well) and the plate incubated for 2 h at room temperature. The plate was washed five times with wash buffer (PBS containing 0.05% (v/v) Tween-20) before being incubated with HRP-conjugated goat anti-mouse IgG antibody diluted 1:5000 in PBS containing 1% (w/v) BSA and 0.1% (v/v) Tween-20 (50 µl/well) for 1 h at room temperature. Following five washes in wash buffer to remove any unbound secondary antibody, 50 µl of 3,3′-5,5′-Tetramethylbenzidine (TMB) (BM blue, POD substrate, Roche Applied Sciences) was added per well and allowed to develop for approximately 30 min in the dark. The reaction was stopped by the addition of 50 µl 0.18 M H_2SO_4 per well. The absorbance at 450 nm was recorded using a plate reader (Multiskan Spectrum, Thermo Fisher Scientific).

2.4. Protein methods

2.4.1. Extraction of total protein from cells

Adherent cells were harvested using a disposable cell scraper and washed once in PBS. Pellets were resuspended in whole cell extract buffer (50 mM Tris pH 8.0, 280 mM NaCl, 0.5% (v/v) NP40, 0.2 mM EDTA, 0.2 mM EGTA, 10% (v/v) glycerol, 0.1 mM sodium vandate; and also 1 mM DTT, 1 mM PMSF added immediately prior to use) and incubated on ice for 10 min. An equal volume of cytoplasmic buffer (20 mM HEPES, 10 mM KCl, 1 mM EDTA, 0.1 mM sodium vandate; and also 1 mM DTT, 1 mM PMSF added immediately prior to use) was added and lysates were

incubated on ice for a further 10 min. Following centrifugation at 16,000 x g at 4^{o}C for 10 min, the soluble extract was removed and transferred to a fresh 1.5 ml tube. Protein was quantified using the Bradford assay (Biorad), measuring the absorbance at 595 nm using a spectrophotometer (6505 UV/Vis Spectrophotometer, Jenway). The concentration was calculated using a standard curve constructed using known amounts of BSA (Roche). Protein extracts were stored at –80^{o}C.

2.4.2. RNase H activity assays

Fluorometric RNase H activity assays were performed using the method described by Crow *et al*. (Crow *et al.,* 2006b) using the RNase H substrate (RNA:DNA hybrid) (forward strand 5′-r(GAU CUG AGC CUG GGA GCU)-fluorescein-3′) and the RNase H2-specific substrate, D-R-D:DNA (forward strand 5′-d(GAT CTG AGC CTG GG)r(A)d(GCT)-fluorescein-3′) where 'd' indicates a deoxyribonucleotide base and 'r' indicates a ribonucleotide base. The forward strands were annealed to the complementary DNA oligonucleotide 5′-AGC TCC CAG GCT CAG ATC-3′which is labelled at the 5′ end with 4-[[4-(Dimethylamino)phenyl]azo]benzoic acid (DABCYL). When the complementary oligonucleotides are annealed together, DABCYL quenches the fluorescent signal. Enzymatic cleavage of the fluorescein-labelled strand releases the fluorescein molecule from the quencher molecule, generating a detectable fluorescence. To calculate the exact amount of fluorescent substrate released in a reaction, a standard curve was prepared by annealing each forward strand with a non-labelled DNA oligonucleotide. Dilutions of 50 pmol, 25 pmol, 12.5 pmol, 3.125 pmol and 0.78 pmol were prepared in assay buffer (60 mM KCl, 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$).

The measurement of RNase H/H2 activity in cell lysates was performed in a 100 µl volume of assay buffer containing 250 nM of substrate and 100 ng/µl of cell lysate (see Section 2.4.1) in a black polystyrene flat-bottomed 96-well plate (non-coated, Costar). Measurement of RNA:DNA hybrid-cleaving activity in preparations of commercially available nucleases was performed using the indicated amount of enzyme and buffer per well (as detailed in Chapter 4).

A set of standards for each substrate was included in every assay. Standards, samples and substrate only controls (to measure background fluorescence from each substrate) were analysed in triplicate. The plate was incubated at 20^{o}C (for cell lysates) or 37^{o}C (for enzymes) and the fluorescent signal read for 100 ms every 5 min for 90 min using a VICTOR^{3} 1420 multilabel counter (Perkin Elmer) with a 480 nm excitation filter and a 535 nm emission filter. A standard curve was constructed and the average pmol of substrate released was calculated for each sample.

2.4.3. Western blotting

30 μg of whole cell lysates from human cell lines were denatured in sample loading buffer (62.5 mM Tris pH 6.8, 50% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.03% (w/v) BPB) for 5 min at 95^{o}C. 100 μg of whole cell lysates from synchronised HeLa cells were pretreated with 2.5 U Benzonase ® nuclease (Novagen) for 15 min on ice, to eliminate any contaminating nucleic acids, before incubation in NuPAGE sample loading buffer (Invitrogen) for 20 min at 60^{o}C. Protein was resolved on pre-cast 4 – 12% NuPAGE® Novex ® Tris-Bis gradient gels (1.5 mm, 10 wells, Invitrogen) using NuPAGE MOPS buffer (Invitrogen). 6 μl of SeeBlue® Plus2 Pre-Stained Standard (Invitrogen) and 4 μl Precision Plus Protein Standard (Bio-Rad Laboratories) were included as size markers. Samples were electrophoresed at a constant 160 V in an XCell *SureLock*^{TM} Mini-Cell electrophoresis system (Invitrogen). Protein was then transferred to a nitrocellulose membrane at a constant 100 V for 1 h at 4^{o}C using a Mini-Trans-Blot Cell system (Bio-Rad Laboratories). Following transfer, membranes were blocked in blocking buffer (5% (w/v) non-fat milk in TBS containing 0.2% (v/v) Tween-20) for 1 h at room temperature. Membranes were probed using a polyclonal sheep antibody raised against recombinant human RNase H2, diluted 1:200 (of a 50% (v/v) glycerol stock) in blocking buffer overnight at 4^{o}C. Following three 5 min washes in washing buffer, the membrane was probed with a HRP-conjugated polyclonal rabbit anti sheep IgG Antibody (Dako Corporation), diluted 1:5000 in blocking buffer for 1 h at room temperature. Unbound secondary antibody was washed away with three 5 min washes in washing buffer and the antigen-antibody complexes were visualised using ECL Western blotting reagents by exposure to photographic film. The membranes

were stripped by incubation in 0.2 M glycine, 1% (w/v) SDS pH 2.5 for 15 min at room temperature. Following three 5 min washes in washing buffer, the membrane was reblocked before being reprobed with housekeeping genes to act as controls for loading. Antibodies used were: polyclonal goat anti-actin (C-11, Santa Cruz Biotechnology) at 1:2000 then donkey anti-goat IgG-HRP (Santa Cruz Biotechnology) at 1:3000; monoclonal anti-α-tubulin (clone B512, Sigma Aldrich) at 1:10,000 followed by goat anti-mouse IgG –HRP 1:5000. Anti-phospho-Histone H3 (Ser-10) (Millipore) diluted 1:5000 detected by HRP-conjugated polyclonal goat anti-rabbit IgG (Cell Signalling Technology) diluted 1:5000 was used as a marker for mitotic cells.

2.5.5. Enzyme-linked immunosorbent assays (ELISAs)

Secreted cytokines in cell supernatants were quantified using sandwich ELISAs. All ELISAs were performed in flat-bottomed 96-well plates (MaxiSorp®, Nunc). Murine IL-6, IL-10, IL-12p40, TNF-α and RANTES were measured using reagents from DuoSet ELISA Development Kits (R&D Systems), according to the manufacturer's instructions. 1% (v/v) Reagent Diluent (R&D Systems) in PBS containing 0.05% (v/v) Tween-20 was used as the buffer for blocking, and for diluting the antibodies, standards and samples.

Murine IFN-α, murine IFN-β and human IFN-α ELISAs were established using the reagents listed in Table 2.2. Each new batch of antibodies and recombinant cytokine standards were optimised for use as there was found to be variation in between different lot numbers. To optimise each new set of reagents, four dilutions each of the capture, detection and standard were tested.

ELISA plates were coated with 50 µl per well of capture antibody diluted in the appropriate buffer and incubated overnight at room temperature. Following five washes in wash buffer (PBS containing 0.05% (v/v) Tween-20), the plates were blocked with 200 µl per well of the appropriate blocking buffer and incubated at room temperature for 1 h. Plates were washed five times in wash buffer and serial dilutions of standards were added in triplicate (50 µl per well). Supernatants were

diluted if necessary in the appropriate buffer and 50 µl per well added to the plate. Each sample was also added in triplicate. Plates were incubated at 4^{o}C overnight to permit capture of the cytokines. Following five washes with wash buffer, 50 µl per well of detection antibody diluted in the appropriate buffer was added and plates incubated for 2 h at room temperature. Unbound antibody was removed by five washes with wash buffer before 50 µl per well of the appropriate HRP-conjugated antibody was added and the plate incubated for 20 min or 1 h (murine IFN ELISAs) at room temperature (except for the human IFN-α ELISA). The plates were washed five times with wash buffer and 50 µl TMB substrate was added to each well, and allowed to develop for 10 – 60 min in the dark. The reaction was stopped with 50 µl 0.18 M H_2SO_4 and the absorbance at 450 nm recorded using a MultiSkan plate reader.

Table 2.2. Reagents used in the ELISAs to detect type I IFNs.

ELISA	Capture antibody	Block	Standard	Detection antibody	HRP-conjugated antibody
Murine IFN-α3	Monoclonal rat anti-mouse IFN- α (Clone RMMA-1), 910 ng/ml (22100-1, PBL Interferon Source)	PBS containing 5% (w/v) BSA	Recombinant murine IFN-α3, 1000 U (12100-1, PBL Interferon Source)	Polyclonal rabbit anti-mouse IFN- α , 80 ng/ml (32100-1, PBL Interferon Source)	HRP-conjugated donkey anti-rabbit, 80 ng/ml (711-036-152, Jackson ImmunoResearch Laboratories)
Murine IFN-β1	Monoclonal rat anti-mouse IFN-β (Clone RMMB-1), 2.5 µg/ml (22400-1, PBL Interferon Source)	PBS containing 10% (v/v) FCS	Recombinant murine IFN-β, 1000 U (12400-1, PBL Interferon Source)	Polyclonal rabbit anti-mouse IFN-β, 500 ng/ml (32400-1, PBL Interferon Source)	HRP-conjugated donkey anti-rabbit, 80 ng/ml (711-036-152, Jackson ImmunoResearch Laboratories
Human IFN-α2	Monoclonal anti-IFN-α, 10 µg/ml (BMS216MST, Bender Medsystems, Human IFN-α matched antibody pairs)	PBS containing 0.5% (w/v) BSA/ 0.05% (v/v) Tween-20	Recombinant human IFN-α2c, 500 pg/ml (BMS216MST, Bender Medsystems, Human IFN-α matched antibody pairs)	Monoclonal anti-human IFN-α HRP-conjugate, 1:1000 (BMS216MST, Bender Medsystems, Human IFN-α matched antibody pairs)	NA

Antibody concentrations given are as a guide only as optimisation of each lot of antibodies was required. All antibodies and samples were diluted in the blocking buffer, except for the Murine IFN-α ELISA, when PBS containing 1% (w/v) BSA was used instead. The top value of the standard protein is given, serial dilutions were made from this to generate a seven point standard curve. NA – not applicable, detection antibody is HRP-conjugated.

2.5.6. Reverse transcriptase activity assay

2 x 10^5 cells of each patient and control LCL were seeded into a 25 cm^3 flask in normal culture medium (Section 2.2.1). Following normal growth for 3 days, 1 ml was taken from the cells and centrifuged at 1500 x g to pellet any cells. The supernatant was transferred to a 1.5 ml tube and stored at –80^oC until it was used for detection of IFN-α by ELISA (Section 2.5.5) or detection of RT activity.

Mg^{2+}-dependent reverse transcriptase activity was measured using a HS-Mg RT activity kit (Cavidi AB), according to the manufacturer's instructions for screening cell supernatants for RT activity. Each well of the 96-well plates provided with the kit was coated with poly (A) RNA, which would act as a template for reverse transcription. Supernatant was incubated with this template in a reaction mixture which also contained a DNA poly (T) primer and BrdUTP. If the supernatant contained RT activity, a DNA strand was synthesised, incorporating BrdUTP, which was detected using an alkaline phosphatase-conjugated antibody against BrdU. Addition of a colourmetric alkaline phosphatase substrate allowed quantification of alkaline phosphatase activity, which was proportional to RT activity. A standard curve was constructed using known concentrations of recombinant HIV-1 RT, therefore the RT activity of a sample was calculated in pg/ml. The detection range of the assay was 0.5 – 500 pg/ml RT activity. Signal was detected using by recording the absorbance at 405 nm using a MultiSkan plate reader, reading the plates at 30 min, 90 min, 120 min and 22 h as recommended by the manufacturer.

2.6. Data analysis

Data was analysed and graphs drawn using Microsoft Office Excel 2007 and Sigma Plot 11.0 (Systat Software, Inc.). Statistical analysis was performed using unpaired Student's *t*-test ($p \leq 0.05$).

Chapter 3. The examination of a potential role for RNase H2 in innate immunity

The cellular functions of the mammalian RNase H2 complex have not been fully characterised. Although hypothesised to be involved in the processing of replication intermediates, its connection with AGS raises the possibility that RNase H2 could function as a component of the innate immune response to nucleic acids, as proposed for TREX1 and SAMHD1 (Rice *et al.*, 2009; Stetson *et al.*, 2008).

RNase H2 is reported to be ubiquitously expressed (Arudchandran *et al.*, 2000; Eder and Walder, 1991; Frank *et al.*, 1998a; Frank *et al.*, 1998c), as is TREX1 (Hoss *et al.*, 1999; Morita *et al.*, 2004). Human SAMHD1 is expressed widely but is absent in thymus and brain tissues (Liao *et al.*, 2008). In this chapter, expression profiling of the five AGS genes was performed to establish in more detail the cell types with high levels of *RNASEH2* transcription, which suggested that RNase H2 may have a specific function in certain immune cells. Together with the published evidence for TREX1 and SAMHD1 as negative regulators of the innate immune response, this led to a series of preliminary experiments to determine if RNase H2 has a role in immunity.

3.1. Human T cells and monocytes contain high levels of active RNase H2

During my MSc maxiproject ('Functional Studies of Human Ribonuclease H2', University of Edinburgh, 2006), I established that expression of the *RNASEH2* genes could be detected in all 22 tissues from a commercially available human tissue panel (Figure 3.1 **a**). This baseline survey of gene expression also revealed that expression of *RNASEH2B* was particularly high in thymus tissue, which raised the initial question: is there a role for RNase H2 in immunity?

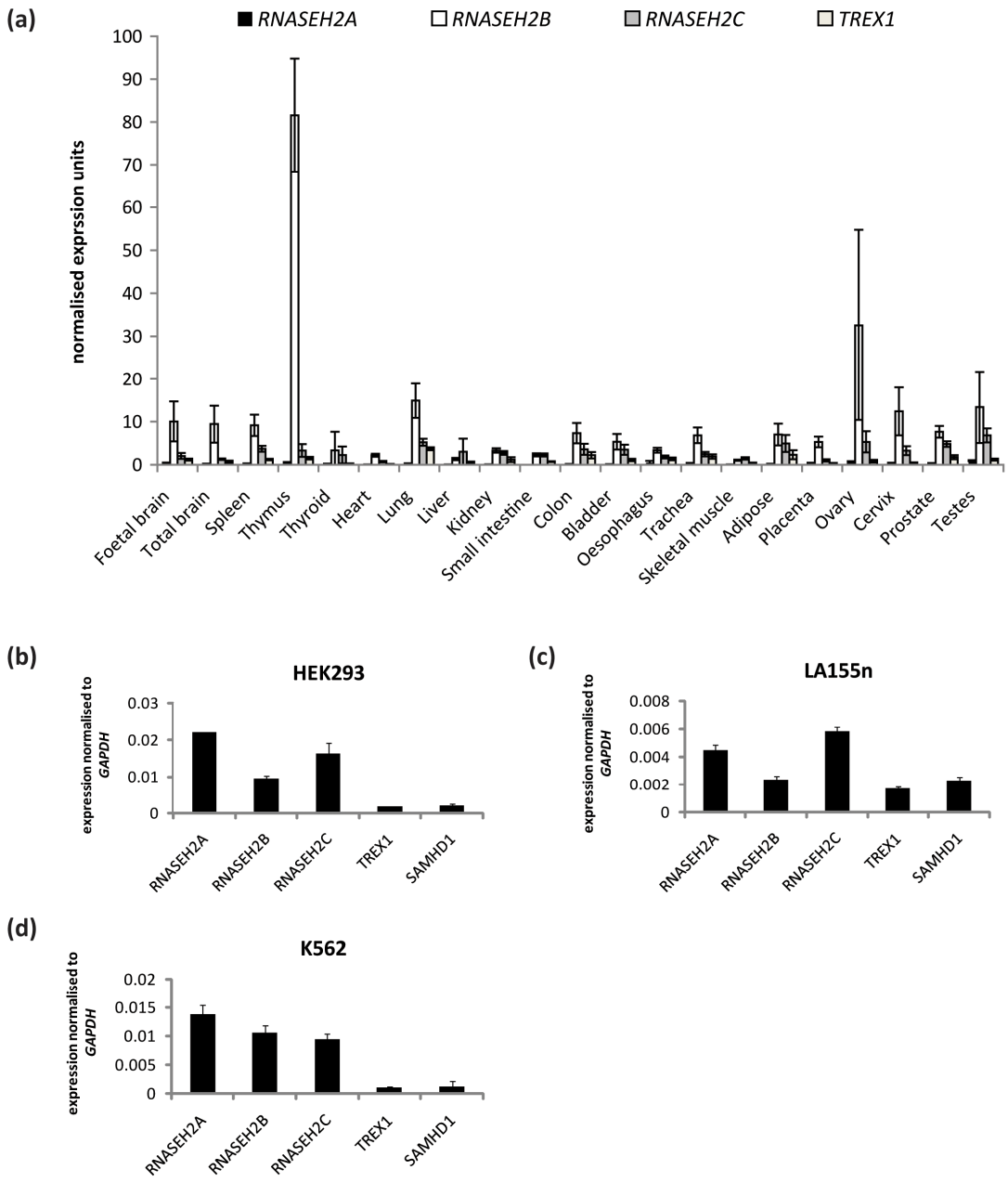


Figure 3.1. Expression profiling of the AGS genes in human cell lines
(a) *RNASEH2* and *TREX1* transcripts can be detected in all tissues in a human tissue panel. Expression was determined by qRT-PCR. Values were normalised to the housekeeping gene *GAPDH*, and the normalised units calculated using a standard curve. *RNASEH2B* expression is approximately 27-fold higher than *RNASEH2C* expression in the thymus. Absolute relative expression levels between tissues can only be compared if it is assumed that levels of the housekeeping gene do not vary significantly between tissues. Data are the shown as mean $\pm$ s.d. of triplicate qPCR wells and figure is reproduced from MSc Maxiproject, 2006. **(b)** The five AGS genes are expressed in an epithelial cell line (HEK293), a neuroblastoma cell line (LA155n) and an erythroleukaemia cell line (K562), as determined by qRT-PCR. Data are shown as the mean expression normalised to *GAPDH* $\pm$ s.d. of triplicate PCR wells, representative of four (K562), three (HEK293) or one (LA155n) independent experiments.

To further characterise the expression patterns of all five AGS genes, endogenous transcript levels were measured in a range of human cell lines. Expression of all five AGS genes could be detected in cell lines of epithelial, neuronal and haemapoetic lineages (Figure 3.1 **b**), which confirmed that the *RNASEH2* genes are widely expressed. In these three cell lines, the normalised expression levels of the *RNASEH2* genes were higher than the levels of *TREX1* and *SAMHD1*.

The expression of the AGS genes in human immune cell lines was also determined. Two T lymphoblastoid cell lines (Jurkat and CCRF-CEM) both had relatively high levels of *RNASEH2B* expression (Figure 3.2 **a**), correlating with the thymus data (Figure 3.1 **a**). In contrast, *TREX1* and *SAMHD1* expression in these cell lines was much lower. This supports previous evidence that *SAMHD1* is not expressed in thymus tissue (Liao *et al.*, 2008). The expression pattern of the AGS genes in B lymphoblastoid cell lines (LCLs) was very different, with relatively high levels of *TREX1* detected in both LCLs in comparison to much lower expression of the *RNASEH2* and *SAMHD1* genes (Figure 3.2 **b**). It has previously been shown that LCLs have high levels of TREX1 activity (Crow *et al.*, 2006a). Together these data demonstrated a difference in the expression of the AGS genes between two different types of lymphocyte, which may be important for the *in vivo* functions of the proteins.

The THP-1 monocytic cell line expressed *RNASEH2B* at a level equivalent to that in Jurkat cells (Figure 3.2 **c**), assuming that expression of the housekeeping gene was constant between cell types. *SAMHD1* expression was also high in THP-1 cells, relative to *RNASEH2A*, *RNASEH2C* and *TREX1* expression. However, when the cells were differentiated into macrophage-like cells using PMA, expression of *RNASEH2B* and *SAMHD1* decreased sharply (Figure 3.2 **c**).

To investigate if the differences in *RNASEH2* expression observed between the different cell lines examined during this survey of human cell types are reflected in the levels of functional protein, total cellular protein extracts were prepared from each of the cell lines. RNase H2 was detected using an antibody raised against all

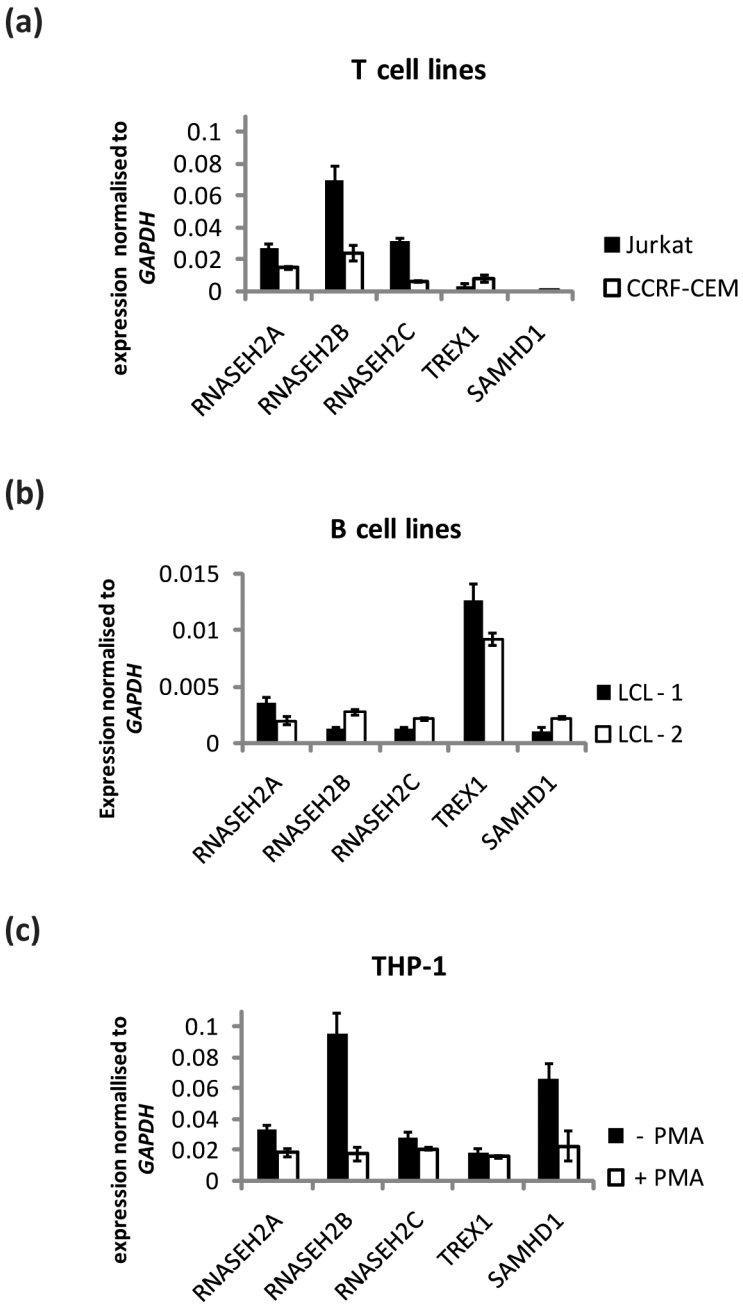


Figure 3.2. Expression of the AGS genes in human lymphocyte and monocyte cell lines. qRT-PCR analysis of AGS gene expression in **(a)** two T lymphoblastoid cell lines, **(b)** two EBV-transformed B lymphpblastoid cell lines derived from two healthy individuals and **(c)** THP-1 monocyte (- PMA) and macrophage-like (+ PMA) cells. Differentiation of THP-1 cells into macrophage-like cells was confirmed by an average 12-fold increase in CD14 expression, as determined by qRT-PCR. Data are shown as normalised mean expression $\pm$ s.d. of PCR wells, representative of four **(a)**, three **(b)** or two **(c)** independent experiments.

three subunits of the complex. Extracts from the two T cell lines contained higher levels of each of the subunits and of the alternative isoform of RNASEH2B than equal quantities of protein extracted from K562 and LCLs (Figure 3.3 **a**). This correlated with a 50% increase in RNase H2 activity in these cells (Figure 3.3 **b**). RNase H2 activity was also increased by 40% in undifferentiated THP-1 cells compared to differentiated cells. This confirmed that in these cell lines, increased *RNASEH2B* expression corresponded to increased RNase H2 enzyme activity.

Together, these data demonstrated that RNase H2 activity was present in a range of different human cell types. The high levels of RNase H2 activity in the two T cell lines and the monocytic cell line may be indicative of a specific function for the enzyme in these cell types.

3.2. The *Rnaseh2* genes are widely expressed in the mouse

As many downstream experiments during this thesis were likely to use murine cells rather than human cells, baseline expression of the *Rnaseh2* genes in murine tissues and cells was characterised using qRT-PCR. This also provided the opportunity to analyse the expression of the genes in primary immune cells, as all human cells used in Section 3.1 were immortalised cell lines.

As shown in Figure 3.4 **a**, expression of the *Rnaseh2* and *Trex1* genes could be detected in adult tissues. Levels of *Rnaseh2b* were particularly high relative to the other genes for each tissue, whereas *Trex1* expression was very low. A similar pattern was seen in primary cells (Figure 3.4 **b**).

Taken together, the preliminary expression data presented here suggests that the *Rnaseh2* genes are widely expressed in mouse, including in cells of the immune system.

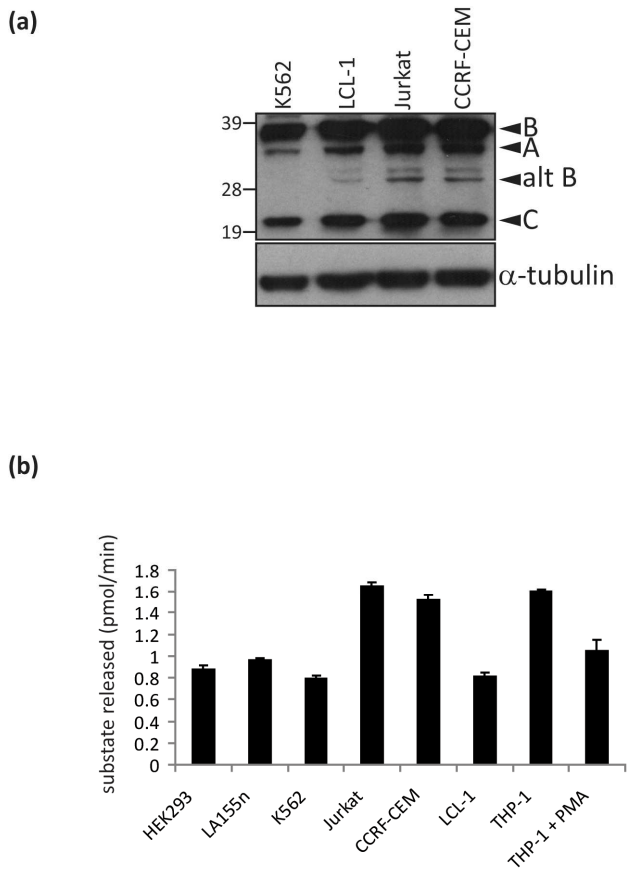
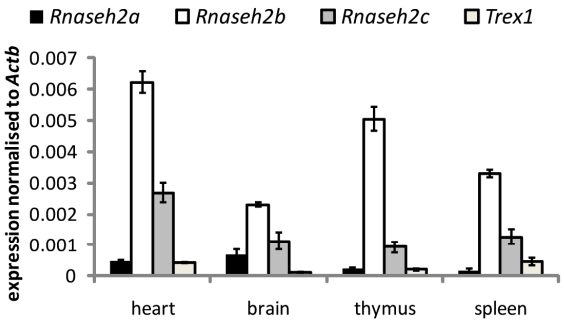


Figure 3.3. Active RNase H2 protein in human cell lines.
(a) Western blot of total cellular extracts from four cell lines to detect the RNASEH2 proteins, including the shorter isoform of RNASEH2B (alt B). 30 µg of whole cell extract were loaded per lane, α-tubulin is shown as a loading control. Blot is representative of two independent experiments. The band observed between A and alt B was due to non-specific binding of the antibody. **(b)** Measurement of RNase H2 activity in total cellular extracts from eight cell lines using an RNase H2-specific substrate. 100 ng/µl of protein was incubated with the substrate at 25°C in a 100 µl volume. The pmol cleaved substrate released over time was recorded for 10 min. Background (substrate only) fluorescence was substracted and RNase H2 activity in pmol/min calculated. Data shown are the mean ± s.d. of triplicate wells, representative of duplicate experiments.

(a)



(b)

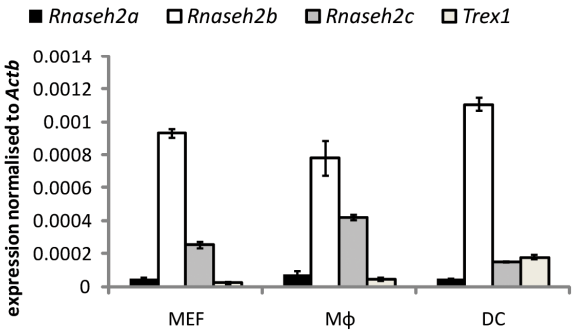


Figure 3.4. Expression of the *Rnaseh2* genes and *Trex1* in the mouse.
qRT-PCR analysis of gene expression in **(a)** adult tissues and **(b)** primary murine embryonic fibroblast (MEF), bone marrow-derived M-CSF macrophage cultures (Mϕ) and bone marrow-derived Flt3-L dendritic cell (DC) cultures. *Samhd1* expression was not examined as these experiments were performed prior to this gene being identified as mutated in AGS. Data shown are mean expression normalised to β-actin (*Actb*) gene expression ± s.d. of triplicate PCR wells. Note that the level of *Actb* expression varied between tissues therefore it is difficult to compare expression levels of the same gene between different tissues but it is possible to compare relative expression levels of different genes within the same sample.

3.3. Levels of RNase H2 are constant throughout the cell cycle

Expression of the *RNASEH2* genes was found to be variable between different tissues. One explanation for this could be the presence of highly proliferating cells in some tissues. In *S. cerevisae*, the mRNA levels of *RNH2A* peak during S-phase and at the G2/M transition, correlating with subsequent peaks in enzyme activity (Arudchandran *et al.*, 2000). To establish if this occurs at the protein level with the mammalian RNase H2 complex, total cellular protein exacts prepared from synchronised HeLa cells were probed with the anti-RNase H2 antibody. Protein levels of the three subunits did not alter throughout the cell cycle (Figure 3.5), which implied that the difference in *Rnaseh2* expression between tissues was unlikely to be attributable to the presence of proliferating cells.

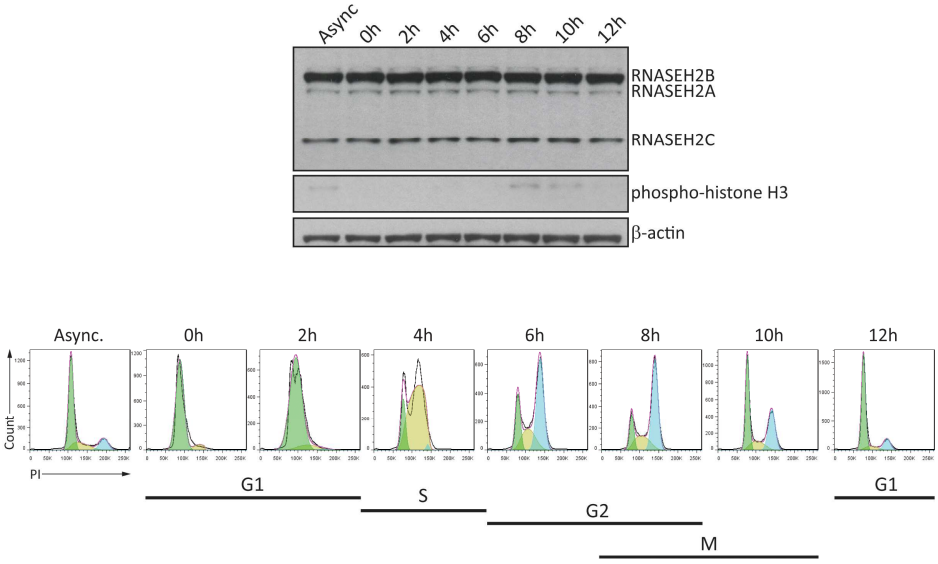


Figure 3.5. Levels of the RNASEH2 proteins do not alter throughout the cell cycle. HeLa cells were synchronised using a double thymidine block before they were released and total cellular protein extracts prepared at the indicated time points by Carol-Anne Martin, MRC HGU. 100 µg per lane of extracts were electrophoresed using a 4 – 12% gradient SDS-polyacrylamide gel before transfer onto a nitrocellulose membrane, which was probed with antibodies against RNase H2, phospho-histone H3 (a marker of mitotic cells) and β-actin (known to remain constant throughout the cell cycle; Carol-Anne Martin, personal communication) as a loading control. Cell cycle stages of the same cells were confirmed using propidium iodide (PI) staining (also performed by Carol-Anne Martin). Cell cycle analysis was performed in FlowJo using the Dean-Jett-Fox model. G1 is shown in green, S phase in yellow and G2 in blue.

3.4. *Rnaseh2* gene expression is not type I interferon-inducible

Expression of *Trex1* and *Samhd1* is induced following the transfection of interferon stimulatory DNA (ISD), in a type I interferon (IFN)-dependent manner (Rice *et al.*, 2009; Stetson *et al.*, 2008). Therefore, to examine the expression of the *Rnaseh2* genes in response to stimulation with this nucleic acid species, primary MEFs were transfected with ISD. This induced a sustained IFN-β response in the MEFs (Figure 3.6 **a**). In agreement with previous experiments performed in macrophages (Rice *et al.*, 2009), *Trex1* and *Samhd1* expression was increased following transfection of ISD (Figure 3.6 **b**). There was a slight increase in expression of *Rnaseh2b* and *Rnaseh2c* after 24 h, however, further replicate experiments would be needed to determine if this increase is statistically significant.

The IFN-β response to ISD in MEFs was established by 4 h (Figure 3.6 **a**). *Trex1* mRNA was upregulated approximately 4-fold after 4 h, but had increased further by 24 h (Figure 3.6 **b**). This suggested that upregulation of gene expression was in response to IFN production and subsequent downstream signalling events, rather than specifically to the ISD itself. To investigate this further, primary MEFs were stimulated with various PRR ligands and expression of the *Rnaseh2* genes and *Trex1* measured.

Following a 24 h stimulation, cells transfected with the dsRNA analogue poly (I:C), a 20 nt Type-A CpG oligodeoxyribonucleotide (CpG ODN) and a 188 nt *in vitro* transcribed (IVT) single-stranded RNA (ssRNA) containing a 5′-triphosphate (5′-ppp) had greatly upregulated expression of type I IFNs (Figure 3.7 **a**). Expression of *Trex1* in these cells was also elevated (Figure 3.7 **b**). In contrast, stimulation with R848 and LPS, which stimulated expression of pro-inflammatory cytokines such as *Il1b* but not type I IFNs, failed to induce *Trex1* expression. Therefore, *Trex1* expression was upregulated in response to type I IFN produced following stimulation of nucleic acid-sensing PRRs of the innate immune system. In contrast, expression of the *Rnaseh2* genes was not induced following stimulation of innate immunity by any PRR ligand tested during these experiments.

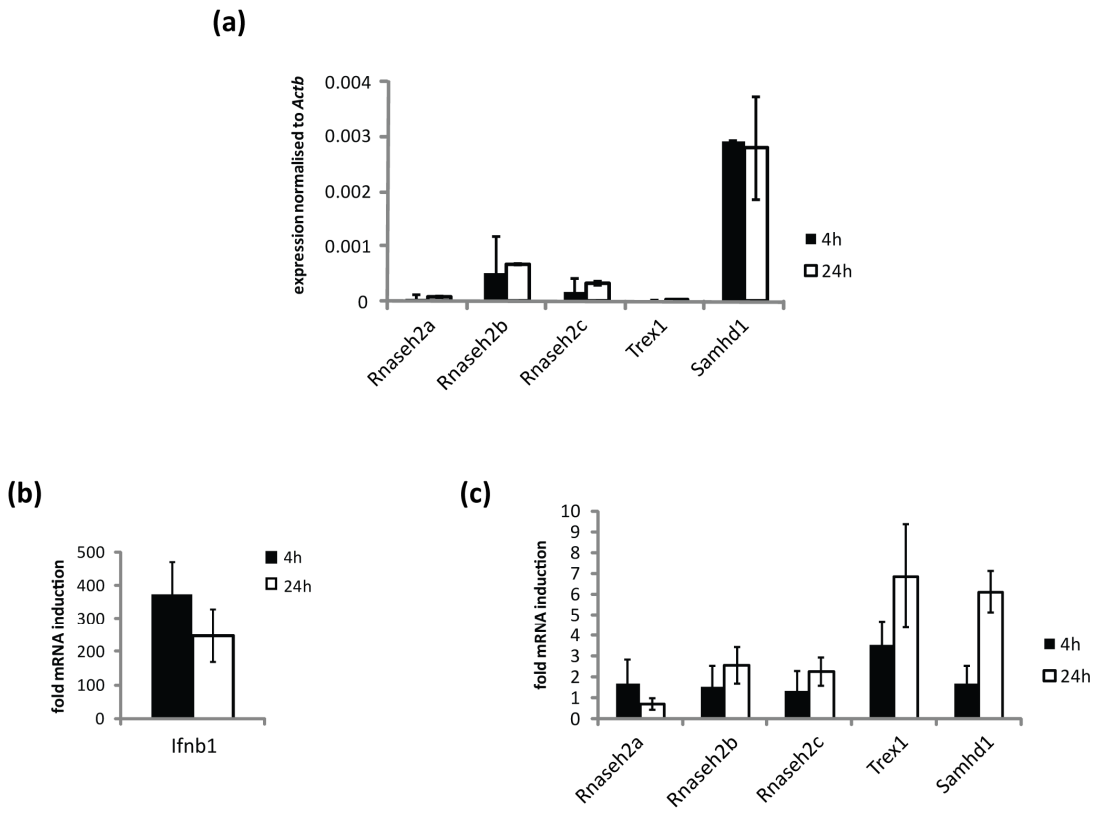


Figure 3.6. Induction of AGS gene expression in MEFs following transfection with ISD. Primary MEFs were transfected with 2 μg/ml of FPLC-purified ISD using Lipofectamine 2000 for 4 h or 24 h. Gene expression was measured using qRT-PCR. **(a)** Basal expression levels of the five AGS genes in primary MEFs. **(b)** Induction of *Ifnb1* expression following transfection of ISD. **(c)** Fold change in expression of the AGS genes in response to ISD. Data shown are represented as normalised (to *Actb* expression) fold mRNA induction from Lipofectamine only controls and are represented as the mean of duplicate experiments ± s.d.

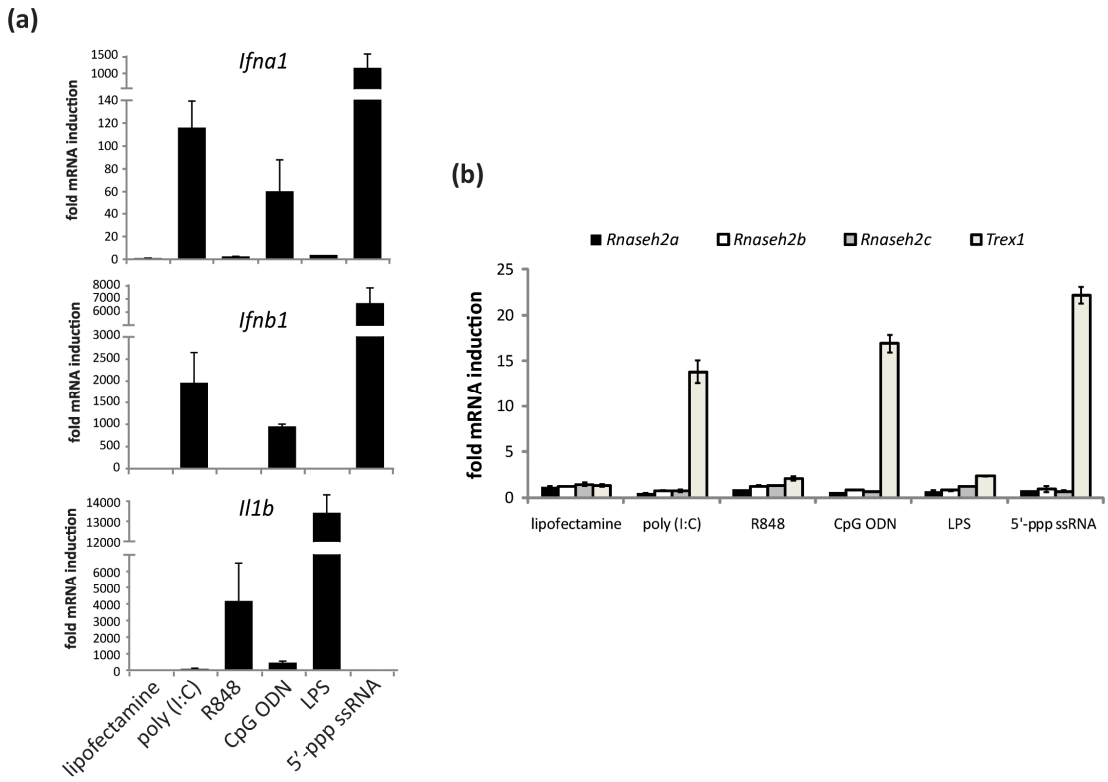


Figure 3.7. Induction of AGS gene expression in MEFs following stimulation with PRR ligands.

Primary MEFs were transfected with 25 μg/ml poly (I:C), 10 μg/ml CpG ODN, 10 μg/ml IVT ssRNA using Lipofectamine 2000. Cells were also stimulated with 2 μg/ml R848 or 200 ng/ml LPS added to the culture medium. After 24 h, cells were harvested and gene expression determined using qRT-PCR. Data shown are fold change relative to medium only controls, and is representative of two independent experiments, ± s.d. of triplicate qRT-PCRs.

To directly establish that *Trex1* and possibly *Samhd1* are upregulated in response to type I IFN but the *Rnaseh2* genes are not, Universal type I IFN was added to the culture medium of primary MEFs and expression of the AGS genes determined over a time course. A rapid increase in the level of *Trex1* mRNA was seen, beginning at only 30 min (Figure 3.8). Induction of *Samhd1* mRNA was later but by 4 h, both *Trex1* and *Samhd1* expression was significantly increased from untreated cells (Student's *t*-test; *Trex1* $p = 0.0006$, *Samhd1* $p = 0.026$). In contrast, even after 24 h, expression of the *Rnaseh2* genes was not significantly altered from 0 h (*Rnaseh2*a $p = 0.157$, *Rnaseh2b* $p = 0.408$, *Rnaseh2c* $p = 0.819$). Overall, this suggests that *Trex1* and *Samhd1* expression is induced in direct response to type I IFN, which is consistent with published data (Rice *et al.*, 2009). In contrast, the *Rnaseh2* genes cannot be induced by type I IFN.

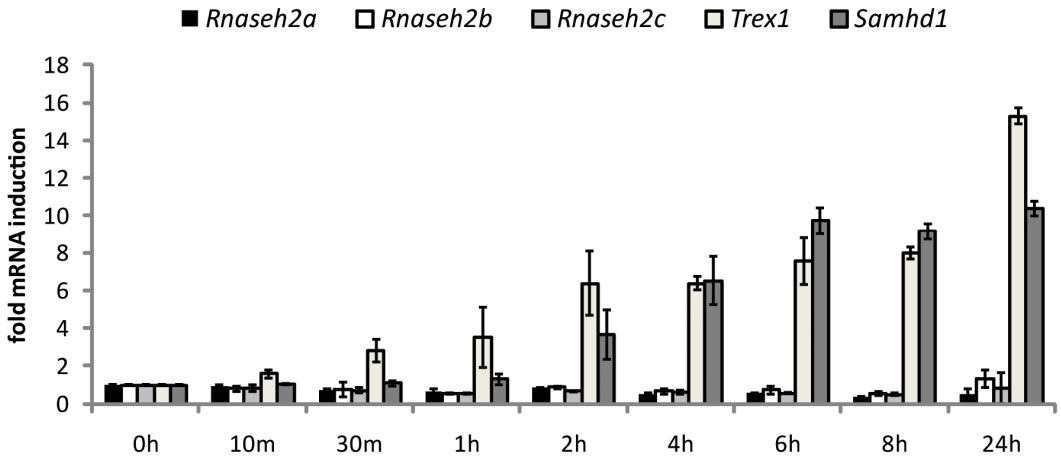


Figure 3.8. Induction of expression of the AGS genes in MEFs following Universal type I IFN treatment.
10^3 U/ml of Universal type I IFN (11200-1, PBL Interferon Source) was added to the culture medium of primary MEFs and the cells lysed for RNA extraction at the indicated time points. Data are shown as the fold mRNA induction from 0 h from two independent experiments using two MEF lines established from separate litters $\pm$ s.d.

In this section, I have demonstrated that the genes encoding the subunits of the RNase H2 complex are not upregulated in response to stimulation of the innate immune response by synthetic PRR ligands, nor as a consequence of IFN-α treatment in primary MEFs. This is in contrast to the other two genes mutated in AGS, *Trex1* and *Samhd1*, which could imply that RNase H2 does not function similarly as a negative regulator of the innate immune response. From the *in vitro* expression data presented in this section, a connection between RNase H2 and the innate immune response remains undetermined.

3.5. The characterisation of RNase H2-deficient cells

The failure of type I IFNs to induce expression of the *Rnaseh2* genes implies that the enzyme complex is unlikely to be a component of an innate immune response pathway. Instead, the function of the enzyme could be to suppress the build up of endogenous nucleic acids, and prevent triggering of cell-intrinsic initiation of autoimmunity by these substrates, which is one role assigned to TREX1 (Stetson *et al.*, 2008). Stetson *et al*. conclude that Trex1-deficiency results in an accumulation of unprocessed nucleic acids, which subsequently trigger an autoimmune response. They show that these nucleic acids are derived from reverse transcribed endogenous retroelements (Stetson *et al.*, 2008). This raises the question of whether RNase H2 functions in the same pathway, potentially in the processing of RNA:DNA hybrid intermediate products of retroelement reverse transcription (Bhoj and Chen, 2008). Therefore, it could be hypothesised that RNase H2-deficient cells will have increased levels of endogenous retroviral replication.

To test this hypothesis, the level of endogenous retroviral replication was characterised in both human and mouse cells deficient in RNase H2 activity.

3.5.1. RNase H2-AGS LCLs do not produce high levels of IFN-α

Several EBV-transformed B lymphoblastoid cell lines (LCLs) from patients with AGS have been established in the lab. LCLs from patients with mutations in the

RNASEH2 genes have a significant reduction in RNase H2 activity compared to control LCLs (Martin Reijns, MRC HGU, personal communication). This could be consistent with an accumulation of nucleic acid substrates which may be derived from endogenous retroelements.

Nucleic acids are potent activators of innate immune signalling, and stimulate a robust type I IFN response (reviewed by Yoneyama and Fujita, 2010). Therefore, the levels of IFN-α in the supernatants of unstimulated RNase H2-AGS LCLs were measured by ELISA. There was no consistent difference between the amount of IFN-α produced by patient LCLs and controls (Figure 3.9). This suggested that IFN-α production was generally not increased in B cells from AGS patients, although this assumed that clonal EBV-transformed B cells are representative of endogenous B cells. Furthermore, this clearly does not exclude the possibility that other cell types from RNase H2-AGS patients may produce significantly elevated levels of IFN-α, particularly cells from the brain (van Heteren *et al.*, 2008). However, due to the very rare occurrence of AGS, the obtaining of other cell types was not feasible during this thesis.

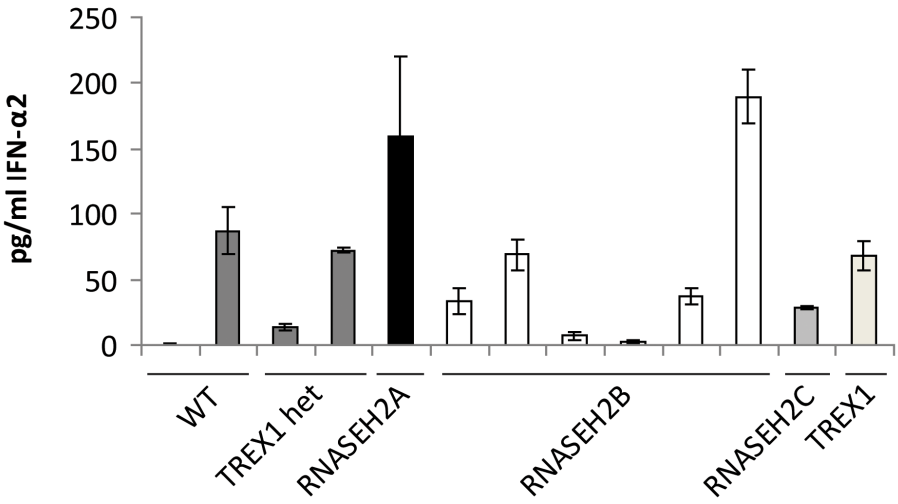


Figure 3.9. Detection of IFN-α in the supernatants of AGS patient LCLs
IFN-α secretion by LCLs was measured by ELISA. Samples were taken from nine AGS patient cell lines with mutations in either the *RNASEH2* genes or *TREX1*. Supernatants from two normal LCLs (WT) and two heterozygous carriers for AGS-causing *TREX1* mutations (TREX1 het) were included as controls. Data were generated from at least two independent supernatant samples per cell line ± s.e.m.

3.5.2. Retroviral replication is not increased in RNase H2-AGS LCLs

To investigate a role for RNase H2 for the suppression of endogenous retroviral replication, the levels of Mg^{2+} dependent RT activity in AGS patient LCL supernatants were measured. This generates a measurement of the retroviral reverse transcriptase (RT) activity of human endogenous retroviruses (HERVs), including the ancestral HERV, HERV-K (Lee and Bieniasz, 2007), and could therefore provide an indication of the production of RNA:DNA hybrid intermediates of reverse transcription in these cells.

The assay was performed using supernatants from AGS patient LCLs with a range of mutations. There was no increase in RT activity in the patient samples (Figure 3.10). All levels detected in the LCL supernatants were greatly below the positive control provided by the manufacturer. Furthermore, qRT-PCR of HERV envelope genes also

suggested there was no notable increase in expression compared to a control LCLs (Appendix 5). Taken together, these data suggest levels of activated endogenous retroviruses are not generally elevated in RNase H2-AGS LCLs. However, this investigation was limited by both the availability of patient material and the types of retroviruses tested for using this assay.

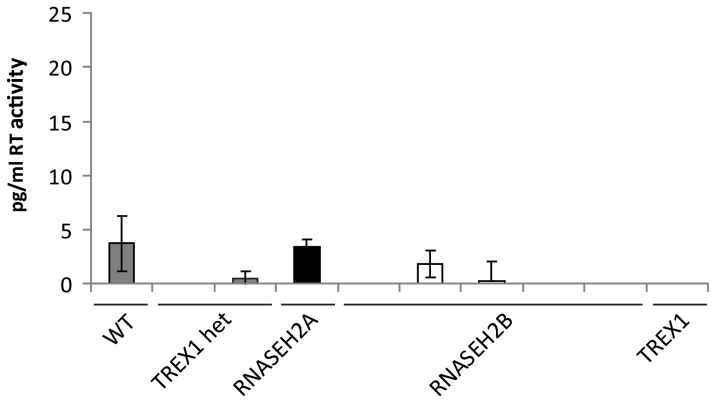


Figure 3.10. Reverse transcriptase activity in supernatants from AGS patient LCLs
Reverse transcriptase (RT) activity was measured in the supernatants of LCL lines as described in Chapter 2 for 22 h. Values were read from a standard curve constructed using recombinant HIV-1 RT and the RT activity of each sample in pg/ml calculated. Data shown are representative of two separate assays using the same supernatants, both performed in triplicate, ± s.d. Previous studies using this assay report samples positive for HERV-K activity as having a minimum of 25 pg/ml of RT activity (Buscher *et al.*, 2005).

3.5.3. Endogenous retroelement expression is not increased in RNase H2-deficient mice

The difficulties associated with obtaining cellular material from AGS patients limited the extent to which RNase H2-deficient human cells can be characterised. Transgenic mice with mutations in the *Rnaseh2b* gene have been established in the lab. To establish if the expression of murine endogenous retroelements was increased in mice with deficiencies in RNase H2 activity, whole-genome gene expression arrays which included probes for retroelement sequences were performed using three of the

transgenic lines. In addition to providing evidence of increased expression of retroelements in RNase H2-deficient cells, analysis of the transcriptomes of the mutant mice was hypothesised to provide further indications as to the *in vivo* function of RNase H2.

3.5.3.1. Rnaseh2bA177T-neo mice

The most frequently observed patient mutation is the exon seven A177T mutation in *RNASEH2B* (Crow *et al.*, 2006b; Rice *et al.*, 2007b). Knock-in mice containing the point mutation and a neo cassette located in between exons six and seven were established, named Rnaseh2bA177T-neo (Figure 3.11 **a**, top). This mutation is hypomorphic: homozygous mice have reduced RNase H2 activity (Figure 3.11 **b**) but are viable and do not develop a disease phenotype (Björn Rabe, MRC HGU, personal communication). MEFs prepared from embryos homozygous for the *Rnaseh2bA177T-neo* allele (n/n) had on average an 87% (Student's *t*-test, $p = 0.0027$) reduction in cellular RNase H activity compared to wild type littermates (Figure 3.11 **b**). RNase H2-specific enzymatic activity is decreased by 84% ($p = 0.0084$). Notably, heterozygous MEFs (n/+) also had a significant reduction in RNase H2 activity, with a 58% ($p = 0.0050$) reduction in RNase H activity and a 61% ($p = 0.015$) reduction in RNase H2-specific activity from wild type MEFs (Figure 3.11 **b**).

Expression microarrays were performed using these primary MEFs lines. RNA was prepared from 3rd passage homozygous mutant MEF lines (n/n) and wild type littermate controls (+/+) from three separate litters. By the 3rd passage, each MEF line should be a homogeneous population of cells which are still actively dividing (Parrinello *et al.*, 2003). The integrity of all RNA used for microarrays was validated using a BioAnalyzer (Figure 3.12 **a**). All RNA samples had a minimum RNA integrity number (RIN) of ≥ 9.0.

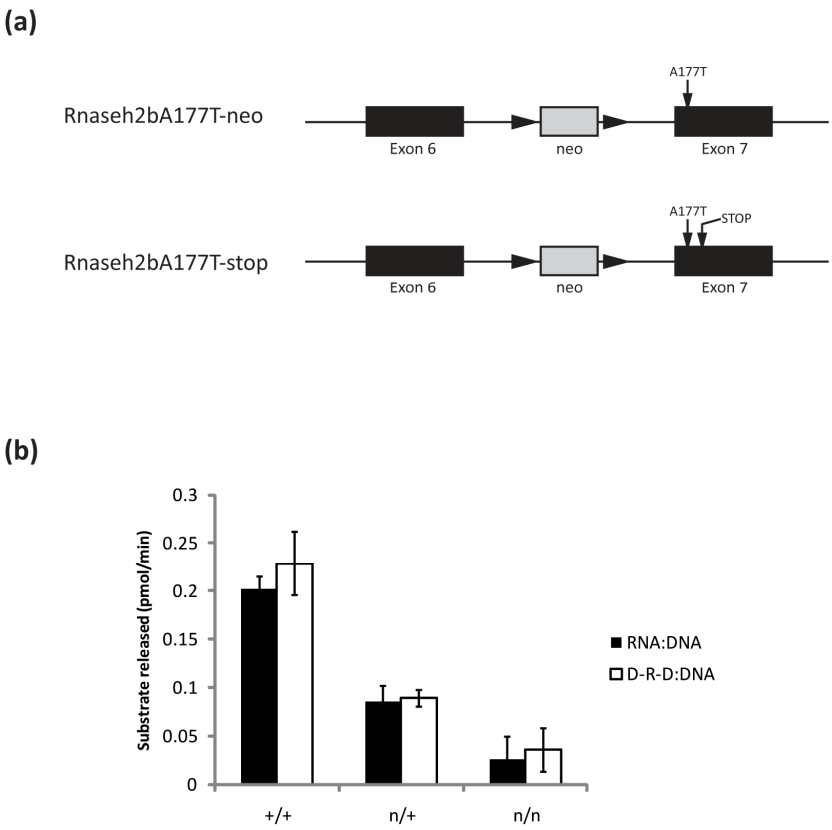
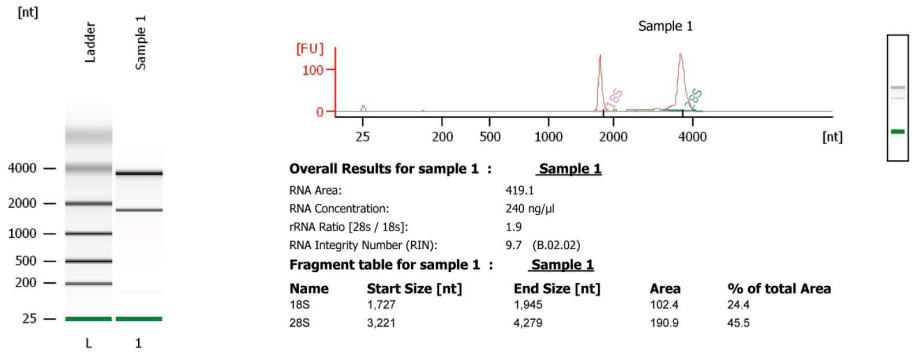


Figure 3.11. Transgenic lines with mutations in *Rnaseh2b*.
(a) Schematic diagram to show the locations of the mutations in the Rnaseh2bA177T-neo and Rnaseh2bA177T-stop transgenic lines. The A177T mutation is a result of a G⟶A missense mutation and the stop is encoded by UGA located 83 bp downstream of this. The neo cassette is flanked by two loxP sites which are denoted by arrowheads. **(b)** Quantification of RNase H activity in MEFs derived from wild type (+/+), heterozygous (n/+) or Rnaseh2A177T-neo homozygous (n/n) embryos. Whole cell lysates were prepared from littermate embryos of each genotype from three separate litters. Cellular RNase H activity was measured using an RNA:DNA hybrid substrate. RNase H2 activity was measured using an RNase H2-specific substrate (D-R-D:DNA). Background fluorescence was subtracted from each data point. Data shown are mean ± s.e.m. (n = 3).

(a)



(b)

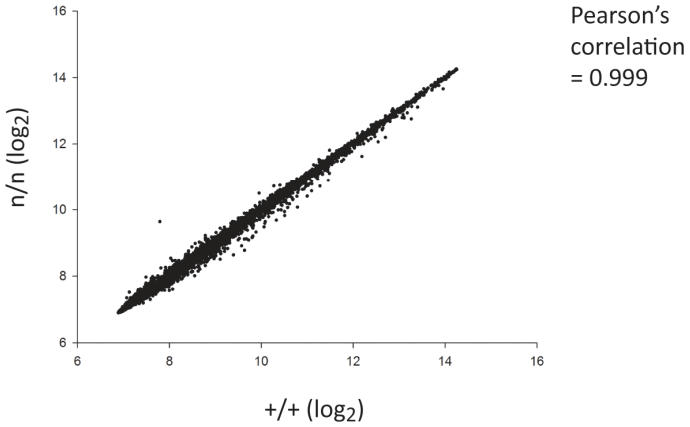


Figure 3.12. Microarray analysis of gene expression in the n/n mouse.
(a) RNA samples were analysed on a RNA 6000 Nano Chip using a BioAnalyzer prior to sample labelling. An RNA Integrity Number (RIN) which is determined by the whole electrophoretic trace, is calculated by the BioAnalyzer. A low RIN (< 7) is indicative of RNA degradation. Data show a representative sample. **(b)** The correlation between the average $\log_2$ expression values for the n/n MEFs versus the +/+ MEFs was assessed using Pearson's correlation measurement, which was calculated using SigmaPlot.

A comparison of the average expression of genes showed that very few were differentially expressed between the three mutant MEF lines and the three wildtype controls (Figure 3.12 **b**). A measure of the correlation of average expression values between the two genotypes using Pearson's correlation gave a value of 0.999, indicating that there is an almost perfectly linear relationship between the two sets of values.

Only 52 transcripts were altered by at least 1.5-fold ($p < 0.05$) in the homozygous mutants; expression of 4 was upregulated and of 48 downregulated (listed in full in Appendix 6), which indicated that there was little difference between the mutant n/n MEFs and wild types at the transcriptional level. The expression of *Rnaseh2b* was significantly reduced in the mutant, providing a positive control for the mutation. Both the microarray probe and the qRT-PCR primers span exons six and seven, therefore this suggests that the transcript is unstable, and may undergo nonsence mediated decay (NMD) to prevent synthesis of a truncated protein. Microarray and qRT-PCR analysis suggest that expression of the other AGS genes is not altered in n/n MEFs (Appendix 7).

There was no significant difference in the expression of endogenous retroelements in the mutants compared to the wild type controls, which reflects the results observed in the AGS patient LCLs. The genes downregulated in the homozygous n/n MEFs did not cluster within any particular functional classification group, and the proportion of genes altered is very small. Therefore, it was difficult to draw any conclusions regarding the *in vivo* consequence of having significantly reduced RNase H2 activity from these data. Consequently, whole-genome gene expression analysis was performed on MEFs from a second mouse line.

3.5.3.2. Rnaseh2bA177T-neo/Rnaseh2bA177T-stop mice

The Rnaseh2bA177T-stop line also contains the A177T point mutation in exon seven downstream of a neo cassette. However, this line also contains a stop codon in exon seven following the A177T transversion (Figure 3.11 **a**). Homozygous Rnaseh2bA177T-stop embryos are not viable (discussed in further detail in Section 3.5.3.3) but Rnaseh2bA177T-neo/Rnaseh2bA177T-stop (n/s) compound heterozygous animals are and do not exhibit a phenotype (Björn Rabe, MRC HGU, personal communication). Therefore, MEFs were established from n/s embryos. RNase H2 activity in n/s cells is reduced by approximately 90% from that of wild type MEFs (Björn Rabe, personal communication).

Whole-genome gene expression microarray analysis of three n/s MEF lines from three separate litters compared to wild type MEFs revealed a large number of genes were differentially expressed; 1257 genes were upregulated (fold change $\geq$ 1.5, $p \leq$ 0.05) and 1829 were downregulated (fold change $\geq$ 1.5, $p \leq 0.05$). Of these genes, 173 were upregulated and 784 down regulated by at least 2 fold in n/s MEFs (top 100 genes are listed in Appendix 8). Expression of *Rnaseh2b* was 2.7 fold lower (p = 0.0000037) in n/s MEFs than in wild type MEFs. The expression of endogenous retroelements were not significantly altered nor were genes involved in innate immune signalling pathways.

Due to time constraints during this thesis, it was not possible to validate the expression data generated from the n/s MEFs using qRT-PCR. However, initial bioinformatic analysis was performed using the web-accessible programs DAVID and FunNet to identify biological pathways putatively affected by the changes in gene expression in n/s MEFs. A high proportion of up-regulated genes were associated with GO terms for apoptosis. Genes associated with protein transport and cell cycle were down-regulated in n/s MEFs (Figure 3.13), which could be indicative of a replication impairment in these cells. However, differential expression of genes associated with the regulation of cell proliferation suggests that n/s MEFs undergo increased cellular proliferation. Therefore, from this bioinformatic overview it is not clear what the consequences of having only 10% of functioning RNase H2 are at the transcriptomic level, which requires further analysis.

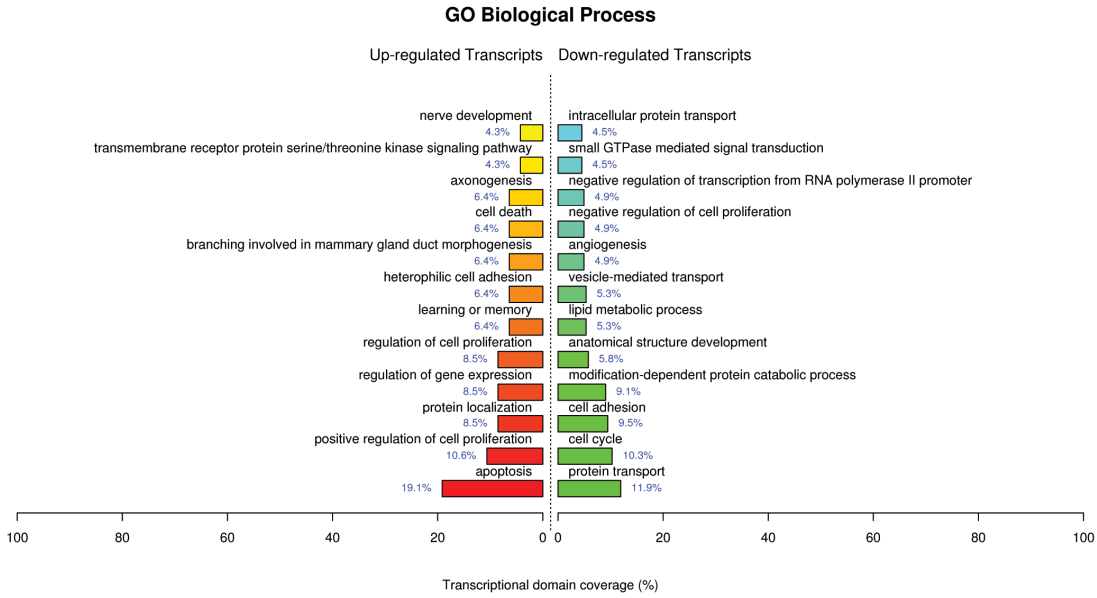


Figure 3.13. Identification of biological processes affected by differentially expressed genes in n/s compound heterozygous MEFs.
Lists of genes differentially expressed in n/s MEFs (fold change ≥ 2, $p \leq 0.05$) were uploaded to the FunNet browser to generate an overview of Gene Ontology (GO) terms for biological processes associated with differentially expressed genes. Data shown here use GO term level 3, reported to provide the optimal balance between the quality and quantity of GO information associated with these genes (Diaz-Uriate *et al.*, 2003).

3.5.3.3. Rnaseh2bA177T-stop mice

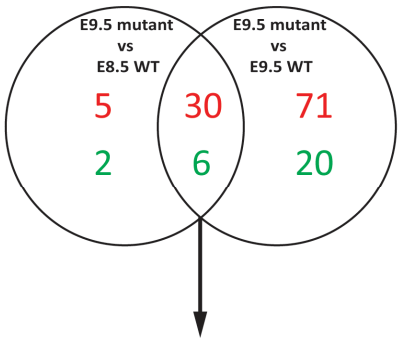
Embryos homozygous for the *Rnaseh2bA177T-stop* allele are not viable. They survive gastrulation (which occurs at E6.5) but fail to develop soon afterwards; by E7.5 they are smaller than wild type litter mates, and by E9.5 a profound developmental delay is evident. Homozygous mutants are reabsorbed by E11.5 (Björn Rabe, MRC HGU, personal communication). Analysis of RNase H2 protein levels in these embryos confirmed that the homozygous embryos are Rnaseh2b-null, and that E9.5 homozygous mutants have an average 92% reduction in RNase H activity and 96% reduction in RNase H2 activity (Björn Rabe, personal communication).

The early embryonic lethality phenotype of this null mutation suggests that RNase H2 is essential for mammalian development. This is in agreement with RNase H2-

AGS patients, who generally have missense hypomorphic mutations in the *RNASEH2* genes (Crow *et al.*, 2006b; Rice *et al.*, 2007b). In contrast, Trex1-null mice are viable (Morita *et al.*, 2004) and TREX1-AGS patients generally have null mutations (Crow *et al.*, 2006a; Rice *et al.*, 2007b). Endogenous retroelements are expressed early in murine development (Baust *et al.*, 2003; Peaston *et al.*, 2004), therefore microarray analysis was performed on homozygous embryos to investigate the expression of these elements compared to wild type embryos. However, it is unclear how an increase in retroelement transcription would lead to developmental arrest.

Whole genome expression microarrays were performed on E9.5 homozygous mutant embryos (s/s) as insufficient amounts of RNA were obtained from earlier stages of development. At E9.5, Rnaseh2b-stop homozygous embryos physically resemble E8.5 wild type embryos (Björn Rabe, MRC HGU personal communication). Therefore, two stages of wild type controls were used for array analysis: E8.5 and E9.5. Three embryos were used for each condition, which were representative of embryos from a number of different litters (Appendix 9).

The E9.5 mutant embryo data were compared to both the E8.5 and E9.5 wild type data sets and lists of differentially expressed genes constructed (Appendix 10). The *Rnaseh2b* gene was downregulated in the mutant embryos by approximately 3-fold compared to both wild-type sets (Figure 3.14). Expression levels of *Rnaseh2a* and *Rnaseh2c* were not significantly altered in the mutant embryos. There was no significant difference in the expression of retroelements between the s/s embryos and the two sets of wild types.



UP

Gene symbol	Gene name	Average fold change v E8.5 WT	Average fold change v E9.5 WT
2010003K15Rik	Hypothetical protein LOC75606	1.92	1.96
4933402E13Rik	Hypothetical protein LOC74437	1.83	1.91
Ass1	Argininosuccinate synthetase 1	2.10	1.59
Bax	BCL2-associated X protein	1.74	1.66
Bbc3	BCL2 binding component 3	2.05	1.82
BC025446	BC025446	1.67	1.66
Ccng1	**Cyclin G1**	**5.86**	**5.27**
Cd81	CD81 antigen	1.56	1.52
Cdkn1a	Cyclin dependent kinase inhibitor 1 (p21)	4.71	4.44
Cldn2	Claudin 2	1.53	1.72
Commd3	COMM domain containing 3	2.30	2.12
Cox6b2	Cytochrome c oxidase subunit VIb polypeptide 2	1.60	2.64
D630023F18Rik	Hypothetical protein LOC98303	1.70	1.58
Ddit4l	DNA-damage-inducible transcript 4-like	2.95	2.37
Def6	Differentially expressed in FDCP 6	1.62	1.68
Got1l1	Glutamic-oxaloacetic transaminase 1-like 1	1.80	1.90
Gpx3	Glutathione peroxidase 3	1.82	1.57
Gtse1	G two S phase expressed protein 1	1.73	1.80
Habp2	Hyaluronic acid binding protein 2	2.13	1.64
Klhl26	Kelch-like 26 (Drosophila)	2.12	1.89
Lgals2	Lectin, galactose-binding, soluble 2	1.69	1.73
Ms4a10	Membrane-spanning 4-domains, subfamily A, member 10	3.55	2.45
Pcbd1	Pterin 4 alpha carbinolamine dehydratase 1	2.14	1.69
Phlda3	Pleckstrin homology-like domain, family A, member 3	4.26	4.05
Slc19a2	Solute carrier family 19 (thiamine transporter), member 2	2.92	2.72
Slc39a4	Solute carrier family 39 (zinc transporter), member 4	2.06	1.63
Svop	Synaptic vesicle 2 related protein	3.22	2.95
Trp53inp1	transformation related protein 53 inducible nuclear protein 1	3.95	3.46
Wig-1	Wild-tye p53-induced gene 1	2.11	2.54
Zfp365	Zinc finger protein 365	2.33	2.22

DOWN

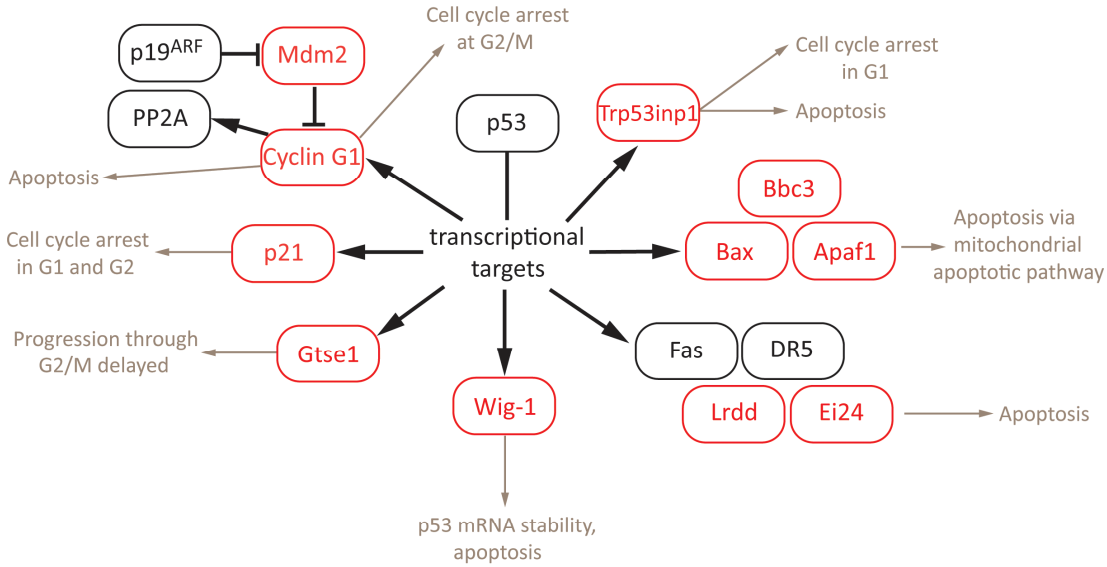
Gene symbol	Gene name	Average fold change v E8.5 WT	Average fold change v E9.5 WT
6330407J23Rik	Hypothetical protein LOC67412	1.75	1.73
Irx3	Iroquois related homeobox 3	1.51	1.59
Lrig1	leucine-rich repeats and immunoglobulin-like domains 1	1.56	1.57
Mbd2	Methyl-CpG binding domain protein 2	1.77	1.65
Rnaseh2b	**Ribonuclease H2, subunit B**	**3.65**	**2.96**
Wnt3a	wingless-related MMTV integration site 3A	1.83	1.59

Figure 3.14. . List of transcripts that are misregulated in the mutant embryos compared to both stages of wild type embryo.
98 genes showed significant ($p \leq 0.05$) changes in expression (> 1.5 fold change), 36 of which were altered in the mutant embryos when compared to both E8.5 and E9.5 wild type embryos. Of the 30 genes significantly upregulated genes, eight are involved in apoptosis/induced in response to DNA damage (shown in blue in the table), including the four genes with the greatest fold increase in expression (*Ccng1*, *Cdkn1a*, *Phlda3* and *Trp53inp1*). Expression of six genes was downregulated in the mutant embryo compare to both wild type sets, but only *Rnaseh2b* expression was decreased by more than 50% (highlighted in bold).

The expression of 35 genes were upregulated in the mutant embryos compared to E8.5 wild type embryos ($p \leq 0.05$, fold change ≥ 1.5), and 101 genes compared to E9.5 wild type embryos (Appendix 10). Thirty of these genes were upregulated in the mutant compared to both wild type sets (Figure 3.14). In turn, eight of these upregulated genes encoded proteins which are involved in apoptosis, p53 signalling and the response to DNA damage, in particular *Ccng1*, *Cdkn1a* and *Phlda3* which were all upregulated in excess of 4 fold in the mutant embryos (Figure 3.15).

Genes which mediate apoptosis and cell cycle arrest are active in normal development, however, the significant upregulation of these genes in the s/s embryos compared to both stages of wild type embryos shows that by E9.5 the mutant embryos are undergoing significant cell death, before they are reabsorbed. Therefore, it is difficult to deduce the *in vivo* function of RNase H2 in development using E9.5 embryos.

In summary, the characterisation of the RNase H2-deficient cells described in this section has not revealed a definitive function for the complex *in vivo*, although it is clear that RNase H2 activity is essential for viability. A significant decrease in RNase H2 activity does not result in elevated expression of endogenous retroelements, therefore the data presented here imply that RNase H2 does not function in the suppression of RNA:DNA hybrids generated by the reverse transcription of endogenous retroelements.



E9.5 mutant vs. E8.5 wildtype

Illumina ID	Gene symbol	Name	Fold increase	P value
ILMN_2920736	Bbc3	BCL2 binding component 3	2.05	3.11E-03
ILMN_2907655	Bax	BCL2-associated X protein	1.74	3.39E-02
ILMN_2634083	Cdkn1a	p21	6.87	3.21E-06
ILMN_2846775	Cdkn1a	p21	4.00	3.43E-06
ILMN_2846776	Cdkn1a	p21	3.25	3.18E-04
ILMN_2500276	Ccng1	Cyclin G1	7.48	3.30E-05
ILMN_2710229	Ccng1	Cyclin G1	5.74	6.01E-04
ILMN_2702233	Ccng1	Cyclin G1	4.35	1.31E-03
ILMN_2506012	Trp53inp1	transformation related protein 53 inducible nuclear protein 1	3.95	6.01E-04

E9.5 mutant vs. E9.5 wildtype

Illumina ID	Gene symbol	Name	Fold increase	P value
ILMN_3143577	Apaf1	Apoptotic protease activating factor 1	1.56	2.52E-02
ILMN_2907655	Bax	Bcl2-associated X protein	1.66	8.72E-03
ILMN_2920736	Bbc3	BCL2 binding component 3	1.82	2.58E-03
ILMN_2500276	Ccng1	Cyclin G1	6.54	1.37E-05
ILMN_2710229	Ccng1	Cyclin G1	5.15	1.89E-04
ILMN_2702233	Ccng1	Cyclin G1	4.12	3.35E-04
ILMN_2634083	Cdkn1a	p21	6.13	4.52E-07
ILMN_2846775	Cdkn1a	p21	4.14	4.52E-07
ILMN_2846776	Cdkn1a	p21	3.05	1.27E-04
ILMN_2953677	Ei24	Etoposide induced 2.4	1.62	8.72E-03
ILMN_2775512	Ei24	Etoposide induced 2.4	1.62	4.38E-02
ILMN_2908065	Gtse1	G two S phase expressed protein 1	1.89	6.67E-03
ILMN_2908070	Gtse1	G two S phase expressed protein 1	1.71	1.49E-03
ILMN_2992009	Lrdd	Leucine-rich and death domain containing	1.60	2.66E-02
ILMN_2850753	Mdm2	Transformed mouse 3T3 cell double minute 2	1.89	3.93E-02
ILMN_2506012	Trp53inp1	Transformation related protein 53 inducible nuclear protein 1	3.46	2.30E-04
ILMN_2418426	Wig-1	Wild-type p53-induced gene 1	2.54	2.17E-04

Figure 3.15. Upregulation of transcriptional targets of p53 in s/s embryos.
A number of the genes that were upregulated in the mutant embryos are targets of p53 (shown in red). The proteins encoded by these genes are involved in a range of cellular processes, predominantly the induction of apoptosis and the arrest of the cell cycle.

3.6. Discussion

The autoinflammatory disease AGS can be caused by loss-of-function mutations in the genes encoding the RNase H2, TREX1 and SAMHD1 proteins (Crow *et al.*,

2006a; Crow *et al.*, 2006b; Rice *et al.*, 2009). In this chapter, I aimed to characterise the expression pattern of the *RNASEH2* genes in immune cells and establish if the enzyme complex has a role in the regulation of innate immunity.

3.6.1. Baseline survey of *RNASEH2* gene expression

Expression analysis confirmed previous observations that the *RNASEH2* genes are widely expressed across a range of tissues and cell types, in both human and murine cells. This also revealed differences in the expression between the AGS genes in different immune cell types, notably in lymphocytes. The *RNASEH2* genes were expressed at relatively low levels in two human B cell lines compared to their expression in T cell lines (assuming expression of the housekeeping gene is constant between the different cell types) (Figure 3.2). Consistent with this, the T cell lines exhibited higher levels of active RNase H2 (Figure 3.3).

The formation of RNA:DNA hybrids at specific regions of the immunoglobulin (Ig) heavy chain gene in B cells is an essential part of class switch recombination (CSR) to switch Ig production (Yu *et al.*, 2003). The low endogenous expression of the *RNASEH2* genes could suggest that RNase H activity is not be required for processing of these R loops, as was originally hypothesised (Yu and Lieber, 2003). However, the two cell lines analysed in Section 3.1 had both been EBV-transformed and therefore may not accurately represent B cells *in vivo* (Siemer *et al.*, 2008). Consequently, it would be of interest to determine the basal expression of the *RNASEH2* genes in primary B cells, and in particular to investigate if they can be induced in response to antigen stimulation, when CSR will be naturally occurring. Defects in CSR have not been studied in AGS patients with loss-of-function mutations in the *RNASEH2* genes, however this could provide further insight into the importance of RNase H2 activity in CSR *in vivo*.

It is not immediately apparent what the *in vivo* requirement for high levels of RNase H2 in T cells and thymic tissue could be. During T cell development within the thymus, different T cell receptors (TCRs) are generated via rearrangement of the four

TCR genes, using V(D)J recombination, in a process similar to that of gene segment rearrangement the Ig genes of B cells (reviewed by Krangel, 2009). However, the formation of RNA:DNA hybrids during V(D)J recombination has not been reported. It could be hypothesised that an RNA:DNA hybrid is generated during TCR rearrangement which requires processing by RNase H2. The detection of high levels of RNase H2 activity in the Jurkat cell line is consistent with this, as TCR rearrangement is still ongoing in these cells, despite them being mature T cells (Zou *et al.*, 2007). It could also be hypothesised that deficiencies in RNase H2 activity lead to abnormalities in thymocyte development. The examination of the TCR repertoire of AGS patients has not been performed, therefore it would be if interest to investigate if any abnormalities are present when RNase H2 activity is reduced. Furthermore, the sequence analysis of V(D)J regions would establish if V(D)J function was normal in these patients.

In addition to T cell lines, high levels of RNase H2 activity were present in the human monocytic cell line THP-1 (Figure 3.2). However, upon differentiation into macrophage-like cells, RNase H2 activity decreased. *RNASEH2* expression was also detected in primary murine bone marrow-derived macrophages and Flt3-L DCs (Figure 3.3). The variation in enzyme activity observed between human immune cell lines suggests RNase H2 could have a specific function in certain immune cell types. Further examination of this may be of interest in defining the *in vivo* function of the enzyme complex.

3.6.2. RNase H2 as a negative regulator of the innate immune response to nucleic acids

TREX1 and SAMHD1 are type I IFN-inducible and are hypothesised to function as negative regulators of the innate immune response to ISD (Rice *et al.*, 2009; Stetson *et al.*, 2008). Therefore, a series of experiments were performed to determine if the RNase H2 complex performs a similar function. It was established that both *TREX1* and *SAMHD1* expression is induced directly in response to type I IFNs in primary MEFs (Figure 3.8). In contrast, expression of the *RNASEH2* genes was not

upregulated in response to stimulation of innate immunity with a range of PRR ligands. This suggests that the *RNASEH2* genes are not inducible by type I IFNs in MEFs and therefore that they are functionally distinct from the other two genes identified as being mutated in AGS. Performing these experiments in additional cell types would establish if this is a general trend. RNase H2 activity may be altered via post-transcriptional regulation of protein levels or modifications, therefore analysis of protein levels and enzyme activity would be required to ascertain if this occurs. Furthermore, analysis of any changes in the subcellular localisation of the complex following type I IFN treatment would determine if relocation of the RNase H2 complex could be of functional significance.

Taken together, the experiments performed in Sections 3.1 to 3.4 of this chapter have demonstrated that whilst RNase H2 is widely expressed, including in at high levels in certain cells of the immune system, transcription of the genes encoding the complex is not induced in response to activation of innate immunity. This is in contrast to TREX1 and SAMHD1, and therefore suggests that the genes mutated in AGS may not directly act in the same biological pathway; more specifically, they are not all acting as negative regulators of the innate immune response.

3.6.3. Expression of endogenous retroelements in RNase H2-deficient cells

The characterisation of cells deficient in TREX1 revealed a role for TREX1 enzymatic activity in the processing of cytosolic DNA. In the absence of TREX1, cells accumulate endogenous nucleic acid species in the cytosol (Stetson *et al.*, 2008; Yang *et al.*, 2007). It is hypothesised that these unprocessed TREX1 substrates are detected by an as yet undetermined cytosolic receptor of the innate immune system, which induces the expression of cytokines, notably the type I IFNs (Stetson *et al.*, 2008). This in turn could drive an adaptive immune response and the subsequent development of autoimmune disease in both humans and mice which lack TREX1 catalytic activity.

It is possible that RNase H2 acts to suppress the accumulation of its substrates, therefore, RNA:DNA hybrids may accumulate in cells with reduced enzymatic acitivity. Inappropriate detection of these RNase H2 substrates by sensors of the innate immune system could lead the induction of a type I IFN response and the establishment of autoimmunity in a manner similar to that in TREX1-deficient individuals. The nature and source of RNase H2 substrates accumulating in the absence of enzymatic activity remains to be established. Reverse transcribed DNA derived from endogenous retroelements has been proposed as one species of TREX1 substrate (Stetson *et al.*, 2008). The reverse transcription process would generate RNA:DNA hybrids, which could normally be processed by the RNase H2 complex. Therefore, it has been hypothesised that cells deficient in RNase H2 activity may have an increased expression of endogenous retroelements, in the absence of the suppression of their replication (Bhoj and Chen, 2008).

The level of endogenous retroviral replication in LCLs established from AGS patients with mutations in the *RNASEH2* and *TREX1* genes was not significantly elevated. Furthermore, whole genome gene expression microarray analysis of mice deficient in RNase H2 activity also failed to detect a significant increase in the expression of endogenous retroelements. Taken together, this implies that RNase H2-deficiency does not result in a global increase in endogenous retroelement expression, and that another cellular source of RNase H2 substrates exists, which remains to be identified.

In summary, the data presented in this chapter demonstrate that the RNase H2 complex is unlikely to be involved in the regulation of the innate immune response via the same mechanism as TREX1 and SAMHD1. The immunological phenotype observed in AGS patients with hypomorphic mutations in the genes encoding the RNase H2 complex, most notably the presence of elevated levels of IFN-α in the CSF and plasma, suggests that activation of the immune system in the absence of infection occurs in these individuals. As nucleic acids are potent stimulators of IFN-α production, it is therefore highly likely that this is due to activation of innate

immunity by nucleic acids. Therefore, it is hypothesised that the endogenous nucleic acid substrates of RNase H2, RNA:DNA hybrids, are responsible for triggering immunity in specific cells. Work is ongoing in the lab to define the nature of nucleic acids which accumulate in cells with reduced RNase H2 activity. However, as the immunostimulatory capacity of RNA:DNA hybrids has not previously been characterised, I aimed to test this *in vitro* by generating synthetic hybrids. Therefore, the next three chapters of this thesis detail the design and characterisation of synthetic RNA:DNA hybrids and their ability to stimulate an immune response *in vitro*. The characterisation of RNase H2-specific substrates that can stimulate the activation of the innate immune system should provide indications as to the endogenous sources of such hybrids, and consequently provide further insight into the *in vivo* function of the RNase H2 enzyme complex.

Chapter 4. Reagents for the study of RNA:DNA hybrids

AGS patients have mutations in genes encoding nucleases, the consequence of which is reduced enzymatic activity (Crow *et al.*, 2006a; Crow *et al.*, 2006b). This is hypothesised to lead to an accumulation of unprocessed endogenous nucleic acid substrates in patient cells. The accumulation of DNA in cells from humans and mice lacking TREX1 activity has been demonstrated (Stetson *et al.*, 2008; Yang *et al.*, 2007). However, as described in the previous chapter, the accumulation of RNA:DNA hybrids in cells with defective RNase H2 activity has not yet been established and work is ongoing in the lab to identify and purify such nucleic acid species.

Many different forms of nucleic acids can activate innate immune signalling via PRRs (Table 1.2). Recognition of endogenous TREX1 substrates is thought to occur via an unidentified cytosolic receptor of the innate immune system, which triggers the induction of type I IFNs in an IRF3-dependent manner (Stetson *et al.*, 2008). However, the ability of RNA:DNA hybrids to stimulate immunity has not been characterised. In addition to identifying a novel nucleic acid ligand for PRRs of the immune system, understanding the mechanism by which RNA:DNA hybrids stimulate immunity may provide further insights into the origin of endogenous hybrids. Reverse transcription of retroviral genomes provide a source of exogenously-derived RNA:DNA hybrids, the immunostimulatory potential of which is unknown. The ability of any of the known PRRs to sense RNA:DNA hybrids has not been described, although RIG-I is able to translocate on a RNA:DNA hybrid in an RNA-dependent manner (Myong *et al.*, 2009).

RNA:DNA hybrids exhibit different physiochemical properties to dsDNA and dsRNA, therefore it was necessary to establish and optimise techniques for performing experiments with hybrids. In this chapter, the unique characteristics of RNA:DNA hybrids are considered. The methodologies for working with hybrids are investigated, including the testing of commercial nuclease preparations for hybrid-

cleaving activity. Finally, applications for an RNA:DNA hybrid-specific antibody are characterised.

4.1. Unique structural and physiochemical properties of RNA:DNA hybrids

4.1.1. The structural properties of nucleic acid duplexes

Studies using various synthetic RNA:DNA hybrids have revealed that their structural conformations are distinct from those of both DNA and RNA duplexes (Noy *et al.*, 2005; Shaw and Arya, 2008). In aqueous solutions, DNA duplexes adopt a B-form structure (B-DNA). Dehydration can induce the DNA to undergo a conformational change into an A-form structure (A-DNA) (Noy *et al.*, 2007). In contrast, RNA duplexes are always A-form, independent of humidity conditions (Figure 4.1). The differences in the helices adopted by RNA and DNA duplexes is considered to be important for their cellular functions, for example, detection by specific proteins and the transfer of biological information (Pan and MacKerell, 2003).

The differences between dsDNA and dsRNA can be attributed to the hydroxyl (OH) group at the 2′ position on the ribose sugar ring, which cannot be accommodated into a B-form conformation because of steric hindrance (Arnott *et al.*, 1986). In addition to lining the minor groove of the RNA duplex, where they are involved in extensive hydrogen bonding with water molecules, intramolecular hydrogen bonds between the 2′-OH also stabilise the RNA duplex (Egli *et al.*, 1996). This contributes to the overall greater thermostability of A-form RNA duplexes compared to both A- and B-form DNA duplexes. The difference in base stacking between A- and B-form helices also confers additional stability to RNA duplexes (Cate *et al.*, 1996).

RNA:DNA hybrids, which contain only one strand of ribose sugars with the 2′-OH, cannot adopt the canonical A- or B-form helices. They instead assume a unique heteromerous (H-form) helix (Arnott *et al.*, 1986; Fedoroff *et al.*, 1993), with physical characteristics intermediate of those for dsRNA and B-form dsDNA (Figure 4.1) (Arnott *et al.*, 1986; Fedoroff *et al.*, 1993; Horton and Finzel, 1996; Hung *et al.*,

1994; Milman *et al.*, 1967; Noy *et al.*, 2005; Wang *et al.*, 1982). This greatly affects the thermostability and electrophoretic migration of the hybrid in relation to dsDNA and dsRNA, and alters the affinity of proteins and compounds to bind them.

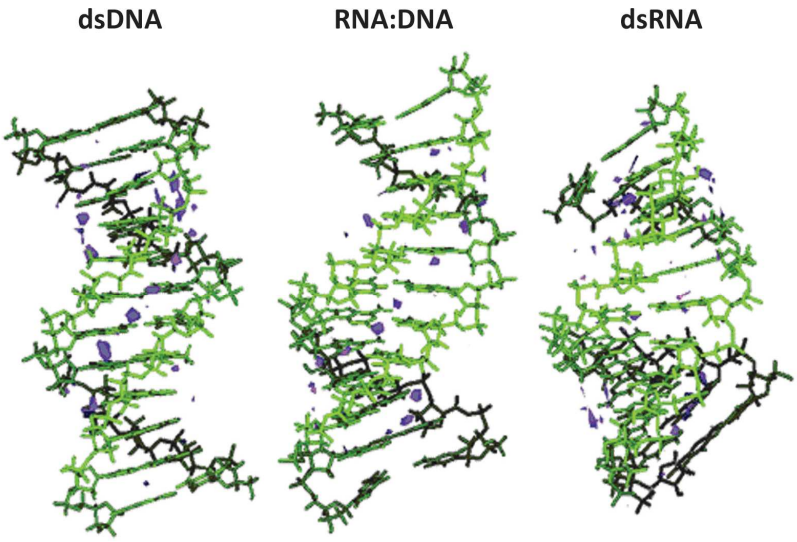


Figure 4.1. Differences in the helical conformations of double stranded nucleic acids. The helical conformation of 12 bp nucleic acids of the same sequence greatly varies depending upon whether the sugars are ribose or deoxyribose. dsDNA adopts a B-form helix whilst dsRNA assumes the shorter, broader A-form. The RNA:DNA hybrid helix is intermediate between the two (H-form), but the measurement of various physical parameters have revealed it is closer to an A-form helix than B-form. (Figure reproduced from Noy *et al.*, 2005).

RNA:DNA hybrids bear a closer physical resemblance to an A-form helix rather than a B-form, but they generally exhibit a reduced thermostability compared to RNA duplexes (Hall and McLaughlin, 1991; Hung *et al.*, 1994). However, this greatly depends on both the base sequence composition of the two strands and the length of the duplex (Hung *et al.*, 1994; Lesnik and Freier, 1995). Analyses of RNA:DNA hybrids consisting of homopurine (Pu) and homopyrimidine (Py) oligonucleotides have concluded that r(Pu):d(Py) hybrids have a greater thermodynamic stability than r(Py):d(Pu) hybrids. Furthermore, the physical structure of the two is very different

(Gyi *et al.*, 1996; Gyi *et al.*, 1998; Hall and McLaughlin, 1991; Hung *et al.*, 1994; Lesnik and Freier, 1995).

The RNA strand of an RNA:DNA hybrid always assumes an A-type conformation, regardless of the sequence (Chou *et al.*, 1989; Priyakumar and Mackerell, 2008; Sanghani and Lavery, 1994). However, the structure of the complementary DNA strand can vary depending on base composition; retaining a B-like conformation when purine-rich but being forced into the adoption of a mixed A-/B-form when pyrimidine-rich and hybridised to a complementary purine rich RNA strand (Gyi *et al.*, 1996; Hung *et al.*, 1994; Noy *et al.*, 2005; Priyakumar and Mackerell, 2008; Salazar *et al.*, 1993). Therefore, r(Pu):d(Py) hybrids are structurally most similar to dsRNA and consequently exhibit a greater thermostability (Gyi *et al.*, 1996; Priyakumar and Mackerell, 2008). Analysis of mixed base oligomeric RNA:DNA hybrids and of naturally occurring hybrid sequences confirmed that individual hybrids form unique helical conformations, which is strongly influenced by the exact sequence composition of the individual strands (Fedoroff *et al.*, 1997; Gray and Ratliff, 1975; Huang *et al.*, 2009; Milman *et al.*, 1967).

Crystallisation of the 3 bp RNA:DNA hybrid r(AAA):d(TTT) flanked by dsDNA revealed that the structure of the central hybrid region is A-form but the DNA duplex regions are B-form (Hsu *et al.*, 2000). Furthermore, the introduction of even a single ribonucleotide into a DNA duplex is sufficient to distort the helix into A-form geometry (Egli *et al.*, 1993). This conformational change is presumably the key to RNase H2 identification and cleavage of such substrates by RNase H2 (Noy *et al.*, 2005).

In summary, the thermostability and structural conformation of RNA:DNA hybrids is greatly influenced by the sequence composition of each strand, which must be given consideration when designing hybrids as potential PRR ligands.

4.1.2. The electrophoretic mobility of RNA:DNA hybrids in polyacrylamide gels

The different physical properties of dsRNA, dsDNA and RNA:DNA hybrids of the same sequence influences their electrophoretic mobility relative to each other. The shorter, broader and less flexible A-form RNA duplex migrates slower through native polyacrylamide gels than the B-form DNA duplex. RNA:DNA hybrid migration is intermediate between the two (Figure 4.2) (Lesnik and Freier, 1995). Therefore, it is possible to distinguish between different nucleic acid duplexes of the same sequence using electrophoresis. Hybrids with a pyrimidine-rich DNA strand migrate more similarly to dsRNA, as their structure is most comparable to A-form (Lesnik and Freier, 1995). As no size markers for RNA:DNA hybrids exist, all markers used throughout this thesis are dsDNA unless stated.

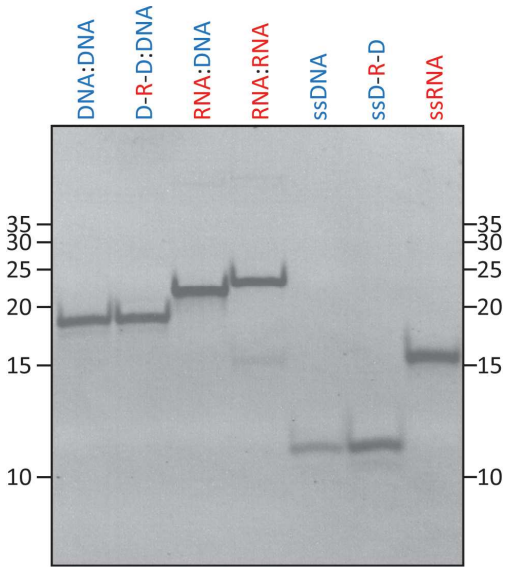


Figure 4.2. The electrophoretic mobility of different nucleic acids.
500 ng of 18 bp duplexes and 18 nt ssNA of the same sequence were electrophoresed on a 20% native polyacrylamide gel. D-R-D indicates DNA containing a single ribonucleotide at position 15. 200 ng of 5 bp ladder was included as a size marker. After staining with acridine orange, the gel was imaged using a UV transilluminator. The Py:Pu ratio of the DNA sequence is 10:8.

A technique for validating the presence of RNA:DNA hybrids in complex nucleic acid mixtures by exploiting their differential migration has been described (Gunnarsson *et al.*, 2006a; Gunnarsson *et al.*, 2006b). Two-dimensional strandedness-dependent electrophoresis (2D-SDE) firstly separates the nucleic acids on the basis of strandedness, length and conformation using polyacrylamide gel electrophoresis (PAGE). The first dimension is performed using a urea-containing gel which reduces the secondary structure of single-stranded nucleic acids (ssNAs) without denaturing double-stranded molecules. After electrophoresis, double-stranded nucleic acids (dsNAs) are heat denatured before samples are separated in the second dimension, which segregates the samples according to length, as they are all now single-stranded. Although this is suitable for identifying RNA:DNA hybrids in mixed nucleic acid samples, it cannot be used to quantify, accurately size or purify the hybrids, therefore was not used during this thesis.

4.1.3. The binding of nucleic acid dyes to RNA:DNA hybrids

The distinct physiochemical properties of RNA:DNA hybrids also affects their binding by chemicals. Studies using the synthetic RNA:DNA hybrids poly (rA:dT) have shown that compounds commonly used to bind the minor groove of dsDNA, such as the fluorescent stains 4',6-diamidino-2-phenylindole (DAPI) and Hoechst, fail to recognise the minor groove of the hybrid (Ren and Chaires, 1999; Wheelhouse and Chaires, 2010). This is likely to be due to the uniqueness of the RNA:DNA hybrid minor groove, which is lined with 2′-OH groups (Horton and Finzel, 1996). In contrast, the intercalator ethidium bromide strongly binds RNA:DNA hybrids (Figure 4.3), and is reported to preferentially bind hybrids over other nucleic acids (Ren and Chaires, 1999; Wheelhouse and Chaires, 2010).

Other compounds which have a high affinity for RNA:DNA hybrids, including ellipticine and propidium, have been described (Ren and Chaires, 1999; Shaw and Arya, 2008), however an RNA:DNA hybrid-specific dye has yet to be identified (Shaw and Arya, 2008).

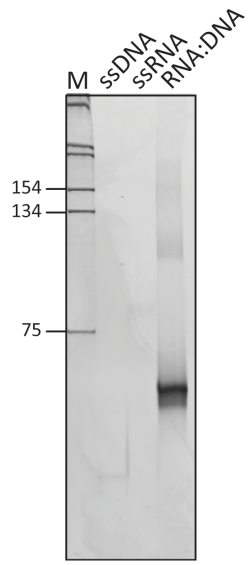


Figure 4.3. Ethidium bromide binds to RNA:DNA hybrids.
A 60 nt oligodeoxyribonucleotide and a 60 nt oligoribonucleotide were annealed to form a 60 bp RNA:DNA hybrid. 500 ng of each was electrophoresed on an 8% native polyacrylamide gel. 200 ng 1kb ladder (M) was included as a size marker. The gel was stained with ethidium bromide and imaged using a phosphorimager.

4.1.4. Quantification of RNA:DNA hybrids

Nucleic acids are typically quantified using two methods. Electrophoresis alongside known amounts of a standard such as a DNA ladder allows comparison of the amount of nucleic acid present in a specific volume. However, as discussed in the previous section, the unique structure of RNA:DNA hybrids can results in the differential binding of fluorescent dyes to them compared to other nucleic acids, specifically dsDNA. Therefore, it would not be possible to compare RNA:DNA hybrids to dsDNA size standards.

UV spectroscopy is used to more accurately quantify nucleic acids. However, the quantification of RNA:DNA hybrids by spectrophotometry has not been well documented. The extinction coefficient for a hybrid is unknown and cannot be calculated simply by adding together the values for the two complementary strands

because of the probable occurrence of hyperchromicity, as known to occur with DNA (Tataurov *et al.*, 2008).

Known molar concentrations of oligonucleotides were annealed to generate 20 μM RNA:DNA hybrids. Electrophoresis on a native polyacrylamde gel confirmed the oligonucleotides had annealed, as evidenced by a shift in electrophoretic migration (Figure 4.4). The molecular weight of each strand was calculated, as described in Chapter 2. This was used to calculate the expected ng/μl of nucleic acid in 1 μl of 20 μM RNA:DNA hybrid. The hybrids were quantified using a NanoDrop UV spectrophotometer and nucleic acid concentrations determined from the absorbance at 260 nm based on the conversion for dsDNA (1 A_{260} unit = 50 μg/ml dsDNA). Each of the single-stranded oligonucleotides was also quantified and the concentration determined using the appropriate conversions from the A_{260} values (1 A_{260} unit = 40 μg/ml ssRNA and 33 μg/ml ssDNA). The expected concentrations for the ssNA were correct to the nearest ng $\pm$ 4.5% (n = 6). For the RNA:DNA hybrids, the recorded concentrations were slightly higher than expected (+ 10% on average), however this could be due to the presence of contaminating unannealed ssNA (Figure 4.4). Nevertheless, this was determined to be the most accurate method of quantifying hybrids, therefore all RNA:DNA hybrids were quantified using the conversion from A_{260} for dsDNA.

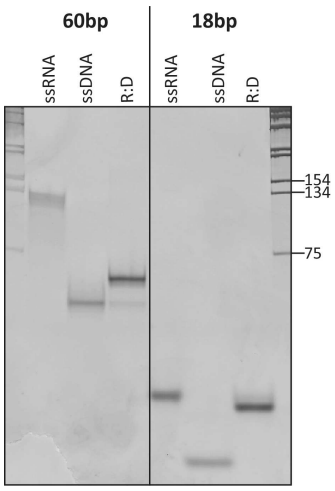


Figure 4.4. Native PAGE analysis of annealed oligonucleotides.
500 ng each (as determined by expected concentration based on MW) of 18 and 60 nt RNA and DNA oligonucleotides and the hybrids formed by annealing the two were electrophoresed on a 6% native polyacrylamide gel, which was stained with acridine orange and imaged using a phosphorimager. 300 ng of a 1 kb ladder was included as a size standard.

4.1.5. Nuclease digestion

RNase H enzymes are the only enzymes to exclusively cleave the RNA strand of RNA:DNA hybrids. However, the multi-functional *E. coli* nuclease Exonuclease III can exonucleolytically degrade both strands of RNA:DNA hybrids (Keller and Crouch, 1972).

Commercial preparations of nucleases are not tested for RNase H activity, and initial experiments performed during this thesis suggested that contaminating RNA:DNA hybrid-cleaving activity was present in many recombinant nuclease preparations. Therefore, the ability of a range of commercially available nucleases to cleave RNA:DNA hybrids was assessed using the RNase H activity assay (Figure 4.5).

4.1.5.1. RNase A will cleave RNA:DNA hybrids at low salt concentrations

The endoribonuclease RNase A cleaves RNA after cytosine and uracil residues to generate 3′ phosphates and 3′ phospho-oligonucleotides products (reviewed by Raines, 1998). It is active under a wide range of conditions; at salt concentrations of less than 100 mM it cleaves ssRNA and dsRNA but at salt concentrations exceeding 300 mM RNase A specifically cleaves ssRNA (Nichols and Yue, 2008; Tanese *et al.*, 1991).

The activity of RNase A on RNA:DNA hybrids in both high (300 mM) and low (60 mM) salt buffers was tested. The amount of enzyme used was determined by the manufacturer's recommended working concentrations for the degradation of RNA. RNase H1 from *E. coli* was included for comparison (Figure 4.5 **a**). The use of high concentrations of enzyme results in cleavage of the RNA:DNA hybrid substrate in the low salt buffer but not in the high salt buffer (Figure 4.5 **b** and **c**). However, ssRNA is still cleaved in the high salt buffer (Figure 4.5 **d**). The activity of RNase A on RNA duplexes in a high salt buffer was found to be highly variable. The ability of RNase A to cleave RNA at such a high salt concentration is unusual as most nucleases, including RNase H and DNase I, are completely inhibited by salt concentrations exceeding 150 mM.

4.1.5.2. RNase I_f will cleave RNA:DNA hybrids at high concentrations

The endonuclease RNase I_f cleaves RNA between all dinucleotide pairs generating products with 5′-OH and 3′ monophosphate ends. According to the manufacturers, it has a strong preference for ssRNA over dsRNA and will not cleave DNA. The ability of RNase I_f to cleave RNA:DNA hybrids was tested using 25 pmol of hybrid substrate. When used at the concentrations recommended for the degradation of 12.5 pmol RNA (12.5 U), RNase I_f exhibited a high level of activity on the hybrid (Figure 4.5 **e**). When used at much lower concentrations, there was still some degradation of the hybrid.

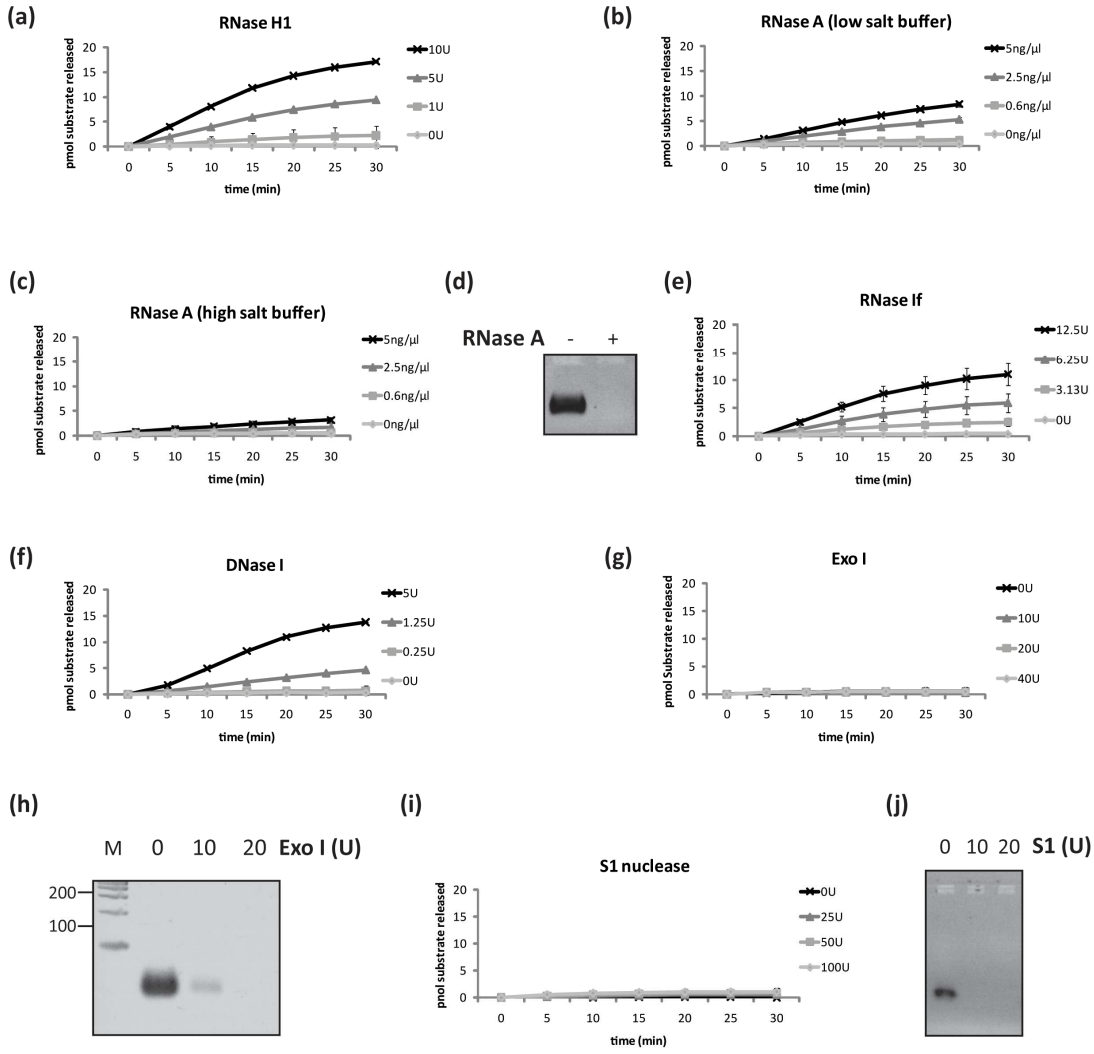


Figure 4.5. The RNA:DNA hybrid-cleaving activity of nucleases.
Commercially available recombinant nucleases were tested for hybrid-cleaving activity using the RNase H activity assay as described in Chapter 2. Different concentrations of nucleases were tested for activity on 25 pmol RNA:DNA hybrid substrate incubated at 37^{o}C. Reactions were performed in a volume of 100 µl. All enzymes units are defined by the manufacturers. **(a)** *E. coli* RNase H (M0297L, New England Biolabs). Reactions were performed in the buffer supplied with the enzyme (75 mM KCl, 50 mM Tris-HCl, 3 mM $MgCl_2$, 10 mM DTT, pH 8.3). **(b)** and **(c)** The activity of DNase-free RNase A (11119915001, Roche Applied Science) was tested in a low salt buffer (60 mM KCl, 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$) and a high salt buffer (300 mM KCl, 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$). **(d)** 2 µg of a 65 nt oligoribonucleotide (93 pmol) was digested with 5 ng/µl RNase A in the high salt buffer. Half of this reaction was electrophoresed on a 1% agarose gel containing ethidium bromide. **(e)** RNase I_f (M0243, New England BioLabs). Reactions were performed in the buffer supplied with the enzyme (10 mM Tris-HCl, 100 mM NaCl 1 mM DTT, 0.5 mM EDTA, pH 8.0).**(f)** RNase-free DNase (M6101, Promega). Reactions were performed in the buffer supplied with the enzyme (40 mM Tris-HCl pH 8.0, 10 mM $MgSO_4$, 1 mM $CaCl_2$). **(g)** *E. coli* Exo I (M0293, New England Biolabs). Reactions were performed in the buffer supplied with the enzyme (67 mM Glycine-KOH, 6.7 mM $MgCl_2$, 10 mM 2-mercaptoethanol, pH 9.5). **(h)** The degradation of ssDNA by Exo I. 1 µg of a 65 nt oligodeoxyribonucleotide (46.5 pmol)

was incubated with the indicated amounts of Exo I. Half of the reaction was electrophoresed on a 2% agarose gel which was then stained with ethidium bromide. **(i)** S1 nuclease (M5761, Promega). Reactions were performed in the buffer supplied with the enzyme (50 mM NaOAc pH 4.5, 280 mM NaCl, 4.5 mM $ZnSO_4$). **(j)** The degradation of ssDNA by S1 nuclease. 200 ng of a 65 nt oligodeoxyribonucleotide (9.3 p mol) was incubated with S1 as indicated. Half of each reaction was electrophoresed on a 2% agarose gel containing ethidium bromide. All enzyme activity experiments are from two independent experiments ± s.d.

4.1.5.3. The cleavage of RNA:DNA hybrids by DNases

To assess the ability of recombinant DNases to cleave RNA:DNA hybrids, DNase I and Exonuclease I (Exo I) were tested. The endonuclease DNase I (here purified from bovine pancreas) cleaves ds- and ssDNA in a sequence-independent manner, resulting in products with 5′-phosphate ends. It does not cleave RNA, however it exhibits weak activity on the DNA strand of RNA:DNA hybrids (Sutton *et al.*, 1997). This was confirmed using the RNase H activity assay; when used at a high concentration, DNase I cleaved the hybrid substrate (Figure 4.5 **f**). However, according to the manufacturer's definition of the activity of one unit of DNase I, 0.25 U should be sufficient to 25 pmol of a DNA substrate. No cleavage of the hybrid was detected when 0.25 U was used (Figure 4.5 **f**), therefore when using DNase I to remove DNA from hybrid containing nucleic acid mixtures, the concentration of enzyme used is crucial to prevent degradation of the hybrid.

The ssDNA-specific *E. coli* Exo I enzyme is a 3′ → 5′ exonuclease. It lacks endonucleolytic activity and cannot act on dsNA. Consequently, even when a great excess of enzyme is used, no cleavage of RNA:DNA hybrids was detected (Figure 4.5 **g**). In contrast, Exo I degraded ssDNA even when used at comparatively low concentrations (Figure 4.5 **h**).

Finally, the hybrid-cleaving ability S1 nuclease was tested. S1 nuclease, purified from *Aspergillus orzyae*, non-specifically cleaves both ssDNA and ssRNA. It has both endo- and exonuclease activities, and releases products with 5′ phosphates. It is approximately five times more active on ssDNA than ssRNA (Vogt, 1973). Although

described as a single-strand specific nuclease, S1 has been shown to introduce breaks into dsDNA, dsRNA and RNA:DNA hybrids when used at high concentrations (Vogt, 1973). However, when tested using the RNase H activity assay, cleavage levels were close to background even at high concentrations (Figure 4.5 **i**). In contrast, use of a low concentration of S1 was sufficient to completely degrade ssDNA (Figure 4.5 **j**).

Together, these RNase H activity assays reveal that caution needs to be taken when using nucleases to degrade unwanted nucleic acids, for example, single-stranded contamination, from preparations of RNA:DNA hybrids.

4.2. The purification and applications of an anti-RNA:DNA hybrid antibody

A number of antibodies with reactivity against RNA:DNA hybrids have been generated (Boguslawski *et al.*, 1986; Kitagawa and Stollar, 1982; Nakazato, 1980). The S9.6 monoclonal antibody is reported to specifically recognise RNA:DNA hybrids in a sequence-independent manner (Boguslawski *et al.*, 1986). The hybridoma line secreting S9.6 was produced from mice immunised with RNA:DNA hybrids generated by the *in vitro* transcription of ΦX174 ssDNA by *E. coli* RNA polymerase (Boguslawski *et al.*, 1986). Initial characterisation of the antibody showed it was specific for RNA:DNA hybrids, and failed to detect dsDNA, ssDNA, dsRNA or ssRNA (Boguslawski *et al.*, 1986).

The S9.6 antibody has been successfully utilised in a range of applications, the majority of which were for the detection of hybrids formed between RNA probes and target ssDNA in an ELISA (Coutlee *et al.*, 1992; Coutlee *et al.*, 1989; Coutlee *et al.*, 1990; Dutrow *et al.*, 2008; Hu *et al.*, 2006; Kinney *et al.*, 1989). However, it has also been used to detect cellular RNA:DNA hybrids by immunofluorescence (Szekvolgyi *et al.*, 2007), electrophoretic mobility shift assays (EMSAs) (Pohjoismaki *et al.*, 2010) and chromatin immunoprecipitation (CHIP) experiments (El Hage *et al.*,

2010). Therefore, the S9.6 provides a powerful reagent for the detection of RNA:DNA hybrids and could potentially be very useful for the validation of hybrids described in this thesis.

The antibody-containing supernatant of the hybridoma cell line HB-8730 (ATCC) was concentrated using a miniPerm system, as described in Chapter 2. The specificity of the antibody was validated by a dotblot using 18 bp RNA:DNA hybrids and controls (Figure 4.6 **a**). The antibody specifically recognised the RNA:DNA hybrid but did not detect the RNase H2-specific substrate (D-R-D:DNA, a DNA duplex containing a single ribonucleotide at position 15 of the forward strand), which suggested that there must be a minimal hybrid region required for binding of the antibody. When used at low dilutions or with longer exposures, faint activity in the supernatant against ssDNA was visible (Figure 4.6 **a**) and a trace of RNase H activity detected (Pohjoismaki *et al.*, 2010). Purification and concentration of the antibody using recombinant protein A/G eliminated all traces of ssDNA reactivity and RNase H activity, which suggested it was due to contaminants in the supernatant rather than the antibody.

The use of the purified S9.6 antibody in techniques that could be useful during this thesis was investigated. One application considered was in the immunodetection of electrophoresed RNA:DNA hybrids. A method for electrophoresing and transferring short (< 200 bp) nucleic acids to a positively-charged nylon membrane was established, as described in Chapter 2. The S9.6 antibody successfully recognised electrophoresed RNA:DNA hybrids (Figure 4.6 **b**). This provided an alternative method to a dotblot for the validation of RNA:DNA hybrid preparations, with the advantage that RNA:DNA hybrids could be accurately sized. As shown in Figure 4.6 **b**, the preparation of 18 bp RNA:DNA hybrid also contained another band migrating at approximately 40 bp. Detection by the S9.6 antibody confirmed that this band was RNA:DNA hybrid-containing. Electrophoresis and Sybr Gold staining of the single-stranded oligonucleotides used to generate this hybrid (ssRNA and ssDNA) revealed that there was an additional band in the ssDNA, migrating at approximately 20 bp (Figure 4.6 **b**). All oligonucleotides used to generate RNA:DNA hybrids were

purified using HPLC or PAGE by the manufacturers. Therefore, this demonstrated that the assumption that commercially purified oligonucleotides are homogeneous in length is not always the case. Consequently, single-stranded oligonucleotides were analysed by denaturing PAGE prior to use to confirm they contained a homogeneous population of nucleic acid.

An additional use for this antibody in the detection of RNA:DNA hybrids in a multiwell plate format was investigated. 18-mer nucleic acids were immobilised using the DNA coating agent Reacti-Bind and an indirect ELISA method with the S9.6 antibody to detect hybrids was established (Figure 4.6 **c**.). The successful use of Reacti-Bind to immobilise DNA to plates has been reported previously (Hirayama *et al.*, 1996). Again, the S9.6 antibody specifically detected the RNA:DNA hybrid. This method could be applied to the detection of RNA:DNA hybrids in synthetic nucleic acid preparations or in the detection and quantification of endogenous hybrids in total nucleic acid extractions from cells.

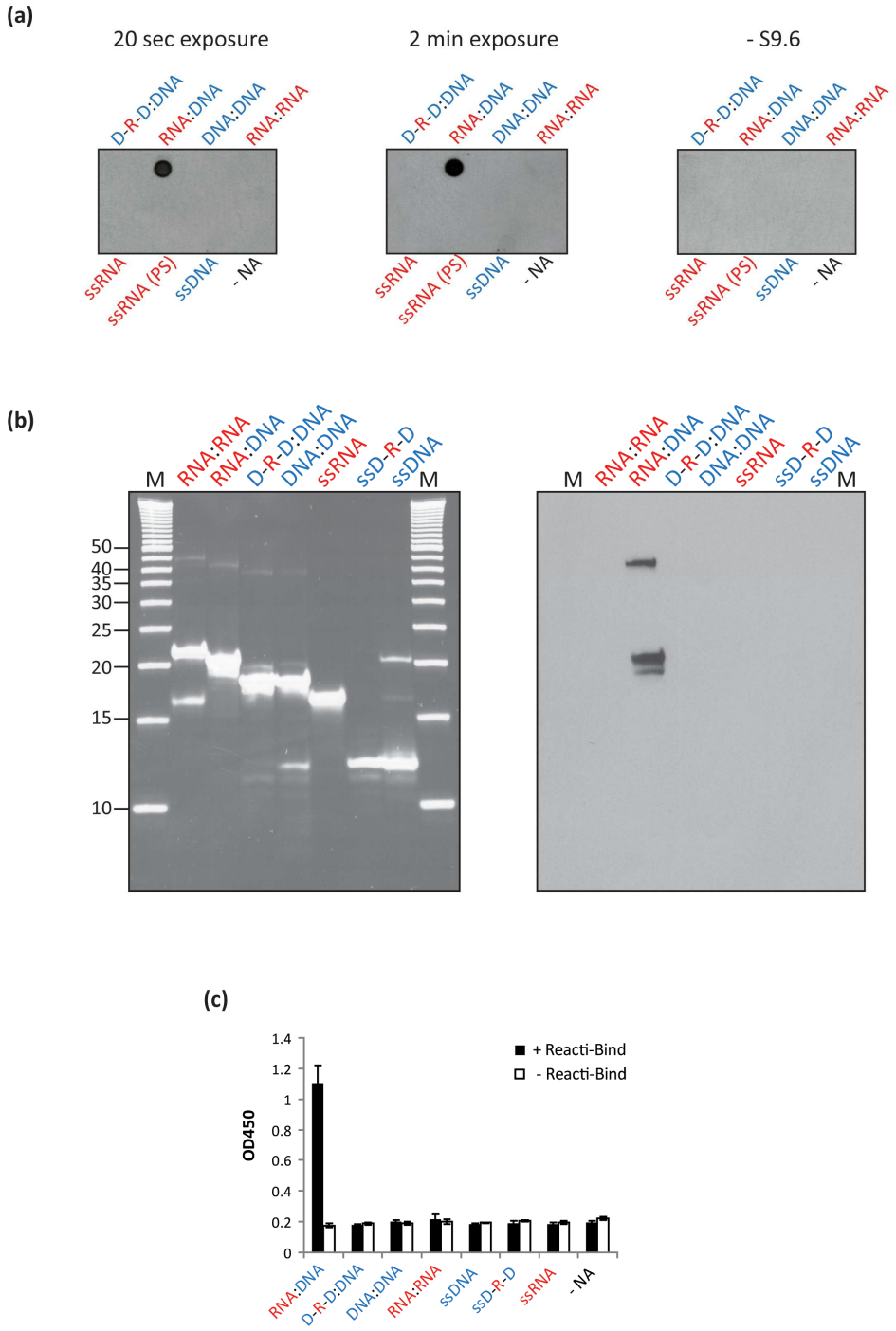


Figure 4.6. Applications of the S9.6 RNA:DNA hybrid-specific antibody.
(a) Characterisation of the specificity of the S9.6 supernatant using 18 nt nucleic acids. ssRNA (PS) is a 20 nt ssRNA oligo with a phosphorothiorate-containing backbone. 20 pmol of each nucleic acid was crosslinked onto a nitrocellulose membrane. The membrane was probed with the concentrated hybridoma supernatant, diluted 1:2000 in PBS containing 2% BSA, as described in Chapter 2. A secondary only (-S9.6) control was included. On the 2 min exposure, faint detection of ssDNA can be seen. **(b)** Immunodetection of electrophoresed RNA:DNA hybrids. 500 ng of each 18 nt nucleic acid was electrophoresed on a 15% native polyacrylamide gel. 500 ng of 5 bp ladder (M) was included as a size marker. The Sybr Gold

stained gel is shown on the left. Following transfer and crosslinking of the nucleic acids to a nylon membrane as described in Chapter 2, the membrane was probed with purified S9.6 in PBS containing 2% BSA. The bands detected by the antibody after a 30 min exposure (right) correspond only to those in the RNA:DNA hybrid lane. The blot is representative of three independent blots. **(c)** Detection of immobilised RNA:DNA hybrids in a multiwell plate. 18-mer nucleic acids at a concentration of 5 ng/µl were immobilised to a 96-well plate in the presence and absence of Reacti-Bind, as described in Chapter 2. Purified S9.6 antibody was used at a 1:1000 dilution in PBS containing 1% BSA. A secondary only control for each nucleic acid was also at background level. Data are representative of triplicate experiments ± s.d. of triplicate wells.

4.3. Discussion

In this chapter, the unique physiochemical characteristics of RNA:DNA hybrids have been considered and their consequences for the study of hybrids discussed. The aim of this chapter was to establish a greater understanding of properties which might affect the design, synthesis and analysis of synthetic RNA:DNA hybrids to be tested for their immunostimulatory potential *in vitro*.

4.3.1. Detection of electrophoresed RNA:DNA hybrids

The physical properties of RNA:DNA hybrids are distinct from those of both RNA and DNA duplexes, most notably in the formation of a unique H-form helical structure (Arnott *et al.*, 1986; Fedoroff *et al.*, 1993). As shown in this chapter, the specific physical properties of hybrids influenced electrophoretic mobility and the binding of fluorescent dyes to the hybrid helix. The differential binding of nucleic acid stains such as ethidium bromide to RNA:DNA hybrids has consequences for both the quantification of hybrids and the analysis of their purity. As shown in Figures 4.4 and 4.6, it is possible that contaminating nucleic acid species may be present in hybrid preparations. The staining of electrophoresed samples using a nucleic acid stain with a high affinity for hybrids would therefore give a misleading representation of the proportion of hybrid compared to contaminating nucleic acids.

Properties distinctive to RNA:DNA hybrids can be exploited. The differential migration of hybrids compared to other nucleic acid species of the same base sequence permits separation of a mixed population of nucleic acids by electrophoresis. Furthermore, the use of the S9.6 antibody to specifically identify hybrids within a mixed population of nucleic acids is also of interest, particularly in the identification of products of annealed oligonucleotides, as shown for an 18 bp hybrid in Figure 4.6. The detection of the extra band at 40 bp by the antibody verifies the presence of an additional hybrid species, which therefore cannot be removed using nuclease treatment. Furthermore, the successful use of the S9.6 antibody to detect electrophoresed nucleic acids also provides a method for detecting RNA:DNA hybrids in RNase H2-deficient cells. Preliminary experiments using a modified version of the method described in this chapter to identify RNA:DNA hybrid regions within electrophoresed total nucleic acid preparations have suggested this is possible, but further refinement of the method would be necessary.

4.3.2. Quantification of RNA:DNA hybrids

RNA:DNA hybrid-specific properties and reagents provide a way of distinguishing them from other nucleic acid species. Equally, the unique characteristics of RNA:DNA hybrids can lead to difficulties, most notably with obtaining an accurate quantification of a hybrid preparation. The absorbance of light at 260 nm differs both between single- and double-stranded nucleic acids and between DNA and RNA. Therefore, the A_{260} of an RNA:DNA hybrid is distinct from that of dsDNA (1 A_{260} unit = 50 μg/ml) and dsRNA (1 $_{A260}$ unit = 40 – 50 μg/ml, (Misquitta *et al.*, 2008)). However, calculating the concentration of a hybrid using the conversion value for dsDNA was found to be sufficiently accurate.

An alternative, highly sensitive approach to the quantification of nucleic acids involves the use of fluorescent dyes, such as PicoGreen® (Molecular Probes, Invitrogen). PicoGreen fluoresces only when bound to dsNA; therefore the precise concentration of a sample can be determined from a standard curve of known dsDNA concentrations (Ahn *et al.*, 1996). Under specific assay conditions, PicoGreen

specifically binds dsNA (Singer *et al.*, 1997). It is widely used for the quantification of dsDNA but has also been used to quantify RNA:DNA hybrids in the same way (EnzChek® Reverse Transcriptase Assay Kit, Invitrogen). However, PicoGreen intercalates at the minor groove of dsDNA (Zipper *et al.*, 2004) and the mechanism of binding to RNA:DNA hybrids has not been shown. Another minor groove intercalator, ethidium bromide, strongly binds electrophoresed RNA:DNA hybrids (Figure 4.3), and is reported to have a higher affinity for binding hybrids over other dsNA (Ren and Chaires, 1999; Wheelhouse and Chaires, 2010). Therefore, it is possible that increased amounts of PicoGreen would bind to RNA:DNA hybrids than would bind to the same concentration of dsDNA. Therefore, the use of a dsDNA standard curve would not accurately determine the concentration of hybrids. Consequently, the use of PicoGreen to quantify hybrids was not pursued during this thesis.

4.3.3. Nuclease digestion of RNA:DNA hybrids

Nuclease digestion provides one method of removing contaminating nucleic acids whilst leaving the RNA:DNA hybrid intact. Commercially available nucleases are not tested for RNase H activity. The two single-strand-specific enzymes, Exo I and S1 nuclease, did not exhibit any activity on the 18 bp RNA:DNA hybrid substrate, even when used at high concentrations, and so could be used in the removal of contaminating ssDNA. In contrast, DNase I exhibited RNA:DNA hybrid-cleaving activity when used at high concentrations. Therefore, a specific amount of enzyme would need to be used to ensure no cleavage of the hybrid, and this would need to be optimised for each hybrid. An additional consequence of the cleavage of the DNA strand of a hybrid by DNase I is the generation of potentially immunostimmulatory ssRNA. Therefore, DNase I is not ideal for use in studies involving RNA:DNA hybrids. Other DNases active against dsDNA including DNase II and Shrimp endonuclease may be suitable; however they would also require testing for hybrid-cleaving activity.

Removal of RNA contamination from hybrid-containing samples using RNase A was also found to be less than straightforward and required optimisation. Low concentrations of enzyme and use of a high salt buffer were required to prevent activity against the hybrid. However, at a high salt buffer RNase A is reportedly also inactive against dsRNA, therefore an additional nuclease would be required if dsRNA contamination was suspected. The dsRNA-specific endonuclease RNase III would not be suitable, as it does not completely degrade dsRNA but instead generates 12 – 15 bp dsRNA cleavage products (Yang *et al.*, 2002a), which again may potentially be immunostimulatory.

In addition to hybrid-cleaving activity when used in low salt buffers, a further disadvantage of utilising RNase A is the cleavage pattern demonstrated by the enzyme on ssRNA. RNase A will cleave only at pyrimidine residues and will not completely degrade ssRNA with a mixed pyrimidine/purine sequence. Instead, short RNA fragments would be produced. In contrast, RNase I_f cleaves ssRNA at every phosphodiester bond; however the nuclease exhibited activity against RNA:DNA hybrids, even when used at low concentrations. It was not possible to determine if this was due to non-specific cleavage by the nuclease or residual RNase H activity from the methods used for commercial production of the recombinant protein. Nevertheless, RNase I_f is therefore unsuitable for the removal of ssRNA contamination from hybrid-containing preparations.

S1 nuclease is reported to be weakly active against ssRNA (Vogt, 1973). It did not show any activity against the 18 bp RNA:DNA hybrid substrate, but is reported to introduce breaks into hybrids when used at high concentrations (Vogt, 1973). Therefore, further experiments would be required to ensure complete degradation of ssRNA without cleavage of the hybrid. The use of nucleases to selectively degrade specific nucleic acids is frequently used as a rapid way of removing contaminating species, however, these enzyme activity assays revealed that the identification of the substrates for each nuclease and optimisation of the reaction conditions would be required to avoid unintentional degradation of the hybrid.

4.3.4. The design of immunostimulatory hybrids

The identification of endogenous RNA:DNA hybrids accumulating in RNase H2-deficient cells is ongoing. This thesis aimed to design and generate synthetic hybrids, therefore an understanding of the unique physical properties of RNA:DNA hybrids was necessary. The sequence of the DNA strand of an RNA:DNA hybrid can have a significant effect on the conformation of the hybrid, as discussed in this chapter. Pyrimidine-rich DNA strands are less resistant to being forced into a mixed A-/B-form by their complementary purine-rich RNA strand. Consequently, these hybrids would more closely resemble dsRNA, which always assume A-form structures. Therefore, by using specific nucleic acid sequences, it would be possible to design hybrids which bear a closer resemblance to dsRNA. The recognition of most nucleic acid ligands by the PRRs of the innate immune system is generally sequence-independent. Instead, characteristics such as length, strandedness and the presence of either a ribose or deoxyribose sugar are key to the sensing of such molecules, as described in Chaper 1, Section 1.5.

The sequences of endogenous RNA:DNA hybrids are largely undetermined, however, R-loops generally contain cytosine-rich DNA strands base paired with a guanosine-rich RNA strand. Therefore, if hybrids accumulating in AGS patients were derived from R-loops, they would structurally resemble dsRNA. Alternatively, the presence of a DNA strand, even a pyrimidine-rich strand, may sufficiently distort the helix to prevent detection by these sensors. In this instance, it would be conceivable that recognition of an RNA:DNA hybrid by a specialised PRR of the innate immune system is strictly dependent on the detection of a unique H-form helix.

In summary, it can be concluded that consideration of the unique physiochemical characteristics of RNA:DNA hybrids is essential when contemplating the use of hybrids for application in downstream experiments. In particular, the validation and optimisation of reagents including nucleic acid gel stains and nucleases is required. The design and synthesis of different RNA:DNA hybrids is therefore described in detail in Chapter 5.

Chapter 5. The synthesis of RNA:DNA hybrids with different physiochemical properties

Many studies characterising innate immune response pathways to nucleic acids utilise synthetic RNA and DNA analogues, such as poly (I:C) and poly (dA:dT), or chemically synthesised oligonucleotides, for example phosphorothiorate modified CpG ODNs (summarised in Table 1.1, Chapter 1). A synthetic RNA:DNA hybrid analogue, poly (rA:dT), is commercially available. This hybrid has been used in studies of hybrid structure (Zimmerman and Pheiffer, 1981), RNase H activity (Nowotny *et al.*, 2005), and is known to be an RNase H2 substrate *in vitro* (Chon *et al.*, 2009).

In addition to using poly (rA:dT), I aimed to design and synthesise RNA:DNA hybrids likely to be immunostimulatory based upon the properties of well characterised nucleic acid ligands for PRRs. *In vitro* testing of hybrids required the generation of sufficient amount of hybrid for multiple experiments and so this was an important consideration when developing methods to synthesise hybrids. In this chapter, the development and optimisation of strategies used to generate hybrids are described.

5.1. The design of RNA:DNA hybrids

5.1.2. Sequence

Innate recognition of viral nucleic acids by PRRs immune system has been extensively studied (reviewed by Yoneyama and Fujita, 2010). Viral RNA in particular potently stimulates type I IFN via TLRs/RLRs in many cell types (as described in Chapter 1, Section 1.5). Therefore, synthetic RNA:DNA hybrids with viral RNA sequences were designed (Table 5.1).

Two short hybrids containing sequence from the HIV-1 LTR were generated. The 18 bp hybrid (R:D18) was already used as the general RNase H substrate in the fluorometric RNase H activity assay developed in the lab (Crow *et al.*, 2006b) and R:D12 is a shortened version of this. The sequence of the 20 nt TLR7 ligand known RNA40 (Heil *et al.*, 2004) is taken from the HIV-1 U5 region. This was used to generate a 20 bp hybrid, termed R:D20(PS). Finally, as GU-rich oligonucleotides and viral sequences are known to stimulate cytokine production through TLR7 (Forsbach *et al.*, 2008; Heil *et al.*, 2004), two hybrids with a repetitive GU motif were generated.

Table 5.1. Properties of the RNA strand of RNA:DNA hybrids designed during this thesis.

Name	RNA sequence	Source	Sugar	Linkage
R:D12	AGCCUGGGAGCU	HIV-1 LTR	2′-O-Me	PD
R:D18	GAUCUGAGCCUGGGAGCU	HIV-1 LTR	Ribose	PD
R:D20(PS)	GCCCGUCUGUUGUGUGACUC	HIV-1 U5	Ribose	PS
R:D20	$(GU)_{10}$		Ribose	PD
R:D60	$(GU)_{30}$		Ribose	PD
R:D188	GGCAAGUACGCCCCCUAUUGACGUCAA UGACGGUAAAUGGCCCGCCUGGCAUUA UGCCCAGUACAUGACCUUAUGGGACUU UCCUACUUGGCAGUACAUCUACGUAUU AGUCAUCGCUAUUACCAUGGUGAUGCG GUUUUGGCAGUACAUCAAUGGGCGUG GAUAGCGGUUUGACUCACGGGGAUUUC	CMV promoter	Ribose	PD

Abbreviations: 2′-O-Me, 2′-O-Methyl ribose; PD, phosphodiester; PS, phosphorothioate

5.1.2. Length

The nucleic acid-sensing PRRs of the innate immune system are specialised for the detection of very different nucleic acid species (Table 1.1, Chapter 1). The membrane-bound receptors TLR7 and TLR9 can both detect short oligonucleotides (20 nt long) (Bauer *et al.*, 2001; Diebold *et al.*, 2004; Heil *et al.*, 2004). RIG-I

signalling can also be stimulated by short RNA oligonucleotides, provided there is sufficient double-stranded region and the presence of a 5′-ppp (Schlee *et al.*, 2009b; Schmidt *et al.*, 2009). In contrast, other cytosolic sensors such as the MDA5 and the AIM2 inflammasome, have a high affinity for longer nucleic acid molecules (Pichlmair *et al.*, 2009; Roberts *et al.*, 2009).

Endogenous RNA:DNA hybrids vary greatly in length. Hybrid regions present in Okazaki fragments are short (8 – 12 nt) (reviewed by Waga and Stillman, 1998). In contrast, R loops formed at the switch regions of Ig genes can exceed 1 kb in size (Yu *et al.*, 2003). The size of pathogenic RNA:DNA hybrids was also considered. Two persistent hybrids which form at the origin of replication of human cytomegalovirus (CMV) are 300 bp and 500 bp in length (Prichard *et al.*, 1998). Therefore, RNA:DNA hybrids ranging in length from 12 bp to 188 bp were designed (Table 5.1).

The most commonly used method of making RNA:DNA hybrids, including for structural (Horton and Finzel, 1996) and biochemical assays (Crow *et al.*, 2006b), is the annealing of an oligoribonucleotide (ORN) with a complementary oligodeoxyribonucleotide (ODN) (Figure 5.1 **a**). Although individual oligonucleotides are often purified by PAGE or high-performance liquid chromatography (HPLC), generally the annealed duplexes do not undergo further purification before use in downstream applications.

This method of hybrid production has many advantages; it is rapid and oligonucleotides can be synthesised with a wide range of backbone and sugar modifications. Therefore it is possible to generate RNA:DNA hybrids with different chemical properties. However, the length of hybrid is limited by the maximum size of ORN that can be efficiently synthesised, which is approximately 80 nt (Sigma-Genosys and Eurogentec). Therefore, an alternative strategy was devised for generating a 188 bp hybrid, which is described in detail in Section 5.3.

5.1.3. Nucleic acid chemistry

In addition to the sequence and length, nucleic acid modifications can affect their immunostimulatory ability (Roberts *et al.*, 2005; Sioud, 2006; Tluk *et al.*, 2009; Wang *et al.*, 2010a). Studies using synthetic nucleic acids to stimulate TLR7 and TLR9 have predominantly used oligonucleotides with a modified chemistry. Phosphorothioate (PS) oligonucleotides, in which one of the non-bridging oxygens in the phosphate group is substituted with sulphur, have an increased thermostability and a greater resistance to nuclease cleavage than oligonucleotides with natural phosphodiester (PD) linkages (Stein *et al.*, 1988).

Single-stranded GU-rich PS-RNA ORNs and PS-containing CpG ODNs are potent TLR7 and TLR9 ligands, respectively (Bauer *et al.*, 2001; Heil *et al.*, 2004). The RNA strand of the hybrid R:D20(PS) was synthesised with a PS-containing backbone, as the stimulation of TLR7 with this sequence was with a PS-modified ORN (Heil *et al.*, 2004). However, all other ORNs used for hybrid production were synthesised with PD-containing backbones (Table 5.1), as were the complementary ODNs. This was chosen because PD-containing hybrids are more physiologically relevant as a PRR ligand than PS-modified hybrids.

ORNs containing modified ribose sugars can also confer a greater stability and nuclease resistance than normal ribose sugars. 2′-O-Methyl ribonucleotides have a 2′-OCH_3 group instead of a 2′-OH and are more resistant to nuclease degradation than normal ribose. An RNA:DNA hybrid with a 2′-O-Methyl RNA strand and a PD-DNA strand is not degraded by RNases H (Crow *et al.*, 2006b). Therefore to prevent potential degradation of the hybrid by cellular enzymes, the RNA strand of R:D12 was synthesised using 2′-O-Methyl RNA.

The RNA:DNA hybrids described in Table 5.1 cover a range of different physiochemical properties, varying in sequence, size and nuclease resistance and therefore provided a starting point for the elucidation of the nature of immunostimulatory RNA:DNA hybrids.

5.2. The generation and purification of hybrids ≤ 60 bp

The short RNA:DNA hybrids detailed in Table 5.1 were formed by annealing equimolar amounts of each strand, heating to 95^{o}C to denature any secondary structure and allowing the mixture to cool to room temperature and form duplexes (Figure 5.1 **a**). Successful annealing was verified by native PAGE analysis, loading a large quantity of nucleic acid to permit detection of any unannealed ssNA (Figure 5.1 **b**).

Three different nucleic acid stains were tested to establish the most appropriate method of detecting single-stranded contaminants using R:D60. As discussed in Chapter 4, ethidium bromide has a particularly high affinity for RNA:DNA hybrids, but as an intercalating dye it does not detect ssNA strongly (Figure 5.1 **b**).

Acridine orange, which intercalates dsNA, can also form electrostatic interactions with ssNA (Traganos *et al.*, 1977). Under UV transillumination at 254 nm, dsNA on PAGE gels stained with acridine orange appear green whilst single-stranded molecules appear red (McMaster and Carmichael, 1977) (as represented in Figure 5. 12). When analysed with a phosphorimager, bands for all three nucleic acid species and the dsDNA ladder are detected (Figure 5.1 **b**).

Acridine orange staining detected only one nucleic acid species in the annealed hybrid, which migrated at the expected size of approximately 60 bp. However, the use of a stain with increased sensitivity revealed the presence of multiple nucleic acid species. Sybr Gold, a highly sensitive nucleic acid stain which binds to both single- and double-stranded DNA and RNA (Tuma *et al.*, 1999), detected two additional bands (Figure 5.1 **b**). One species, indicated by the blue arrow, is unannealed ssDNA60. However, the species which migrated at approximately 100 bp (indicated by the black arrow) was unlikely to be ssRNA60, which co-migrates with the DNA marker at approximately 80 bp. The band at 100 bp was also visible using ethidium bromide staining, which suggested that it contained double-stranded, potentially hybrid regions. RNase H1 digestion confirmed that it was, at least in part, an alternatively annealed RNA:DNA hybrid species (Figure 5.1 **c**).

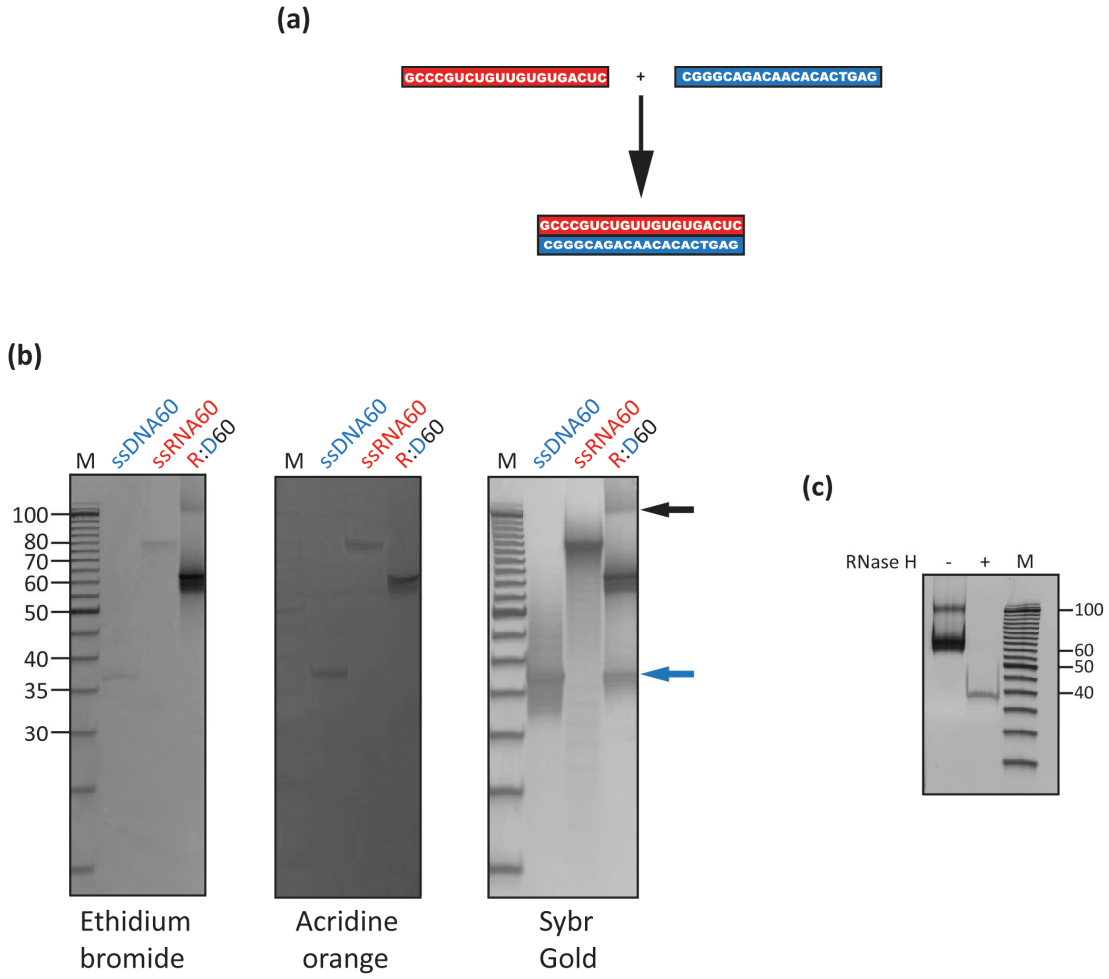


Figure 5.1. Generating short (< 60bp) RNA:DNA hybrids by annealing of oligonucleotides. **(a)** Schematic of hybridisation process, as shown for R:D20(PS). **(b)** PAGE analysis of annealed R:D60. 500 ng each of the single-stranded oligonucleotides (ssDNA60 and ssRNA60) and 500 ng of annealed R:D60 were electrophoresed on 15% native polyacrylamide gel. 300 ng of 5 bp dsDNA ladder (M) was included as a size control. Each section was stained with a different nucleic acid stain, as described in Chapter 2, and imaged using a phosphorimager (excitation at 532 nm, with SHG laser (green)). Blue arrow indicates ssDNA contamination, black arrow indicates an additional nucleic acid species. **(c)** 500 ng of annealed R:D60 was incubated in the presence or absence of 10 U of *E. coli* RNase H1 in a volume of 20 µl containing 1 X reaction buffer supplied with the enzyme. After a 1 h incubation at 37°C, 10 µl of each was electrophoresed on a 15% native polyacrylamide gel, with 300 ng 5 bp ladder (M). The gel was stained with Sybr Gold and imaged as described for **(b)**.

Therefore, it was possible to generate large quantities of RNA:DNA hybrids less than 60 bp in length by annealing complementary oligonucleotides. However, PAGE analysis of the annealing products using a highly sensitive nucleic acid stain revealed

the presence of additional nucleic acid species, including ssNA. Some PRRs of the innate immune system are known to detect ssNA (summarised in Table 1.1, Chapter 1). Therefore, in order to determine if immunity can be activated specifically by RNA:DNA hybrids, it was necessary to purify the correct hybrid away from contaminants prior to use in cell stimulation experiments.

5.2.1. Purification of annealed RNA:DNA hybrids

Initally, RNA:DNA hybrids were PAGE purified. Analysis of gel extracted hybrids showed that they were successfully purified from contaminating ssNA; however the yields recovered were very low (generally less than 30% of input), making this an impractical method for generating large batches of hybrid for subsequent experiments. In addition, contaminants from the process, primarily water-soluble acrylamide oligomers which can bind to nucleic acid, are difficult to remove from gel purified samples (Lukavsky and Puglisi, 2004). Therefore, an alternative to PAGE purification was established.

5.2.1.1. Fast performance liquid chromatography (FPLC)

Liquid chromatography, most often using a high-performance liquid chromatography (HPLC) system, is routinely employed to purify newly synthesised full-length oligonucleotides from any by-products of synthesis. Size exclusion chromatography using a fast-performance liquid chromatography (FPLC) system to successfully purify homogeneous RNA samples has been described (Kim *et al.*, 2007). As HPLC and FPLC systems are very similar (differing mainly in the maximal pressures to be used with the columns), and an ÄKTA FPLC system was available in the lab, the use of FPLC to separate RNA:DNA hybrids from unannealed single-stranded contaminants was investigated.

5.2.1.2. Size exclusion FPLC can fractionate single- and double-stranded nucleic acids of the same length

According to the manufacturer's instructions, the 24 ml Superdex 200 10/300 GL column is appropriate for the gel filtration of dsDNA molecules less than 200 bp in

length. Nucleic acids behave very differently to proteins under native conditions and can often form extensive secondary structures, in particular ssRNA. Therefore, they fractionate at very different molecular weights (MWs) to proteins (Kim *et al.*, 2007). Superdex 200 10/300 GL columns are routinely used for size exclusion chromatography of proteins, therefore it was necessary to establish the suitability of these columns for the fractionation of a mixture of single- and double-stranded nucleic acids of the same length.

Firstly, the void volume of the column, which is the elution volume of molecules which are too large to enter the pores of the gel, was determined to be approximately 7.5 ml (Figure 5.2). Therefore, in subsequent experiments any nucleic acid eluting in the first 7.5 ml will have passed straight through the column without entering the gel bed.

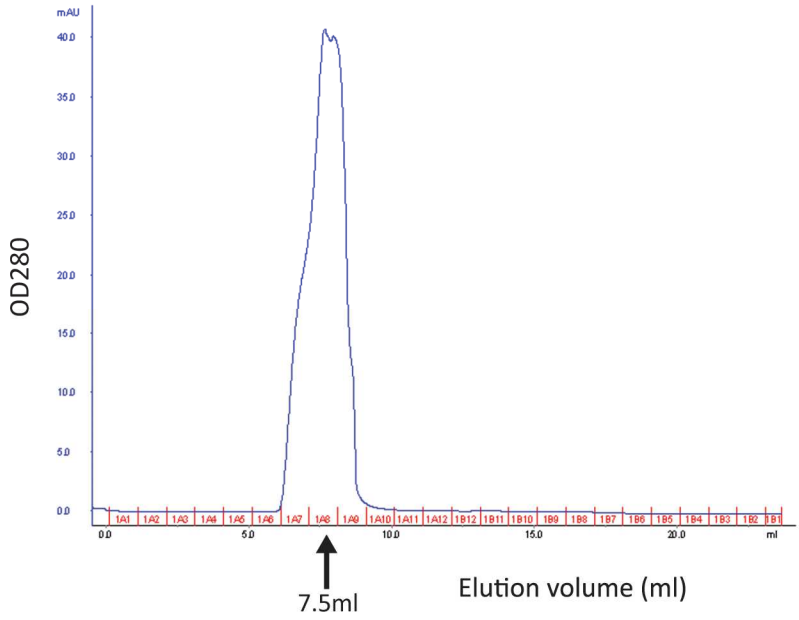


Figure 5.2. Determination of the void volume of a Superdex 200 10/300 GL column
The void volume was determined using 100 µl blue dextran (2000 kDa) from the GE Healthcare calibration kit, eluting with water. The void volume begins at 6 ml and peaks at 7.5 ml.

A 40 nt ODN containing a single ribonucleotide at position 21 was annealed to a complementary ODN. Equimolar amounts of the duplex and the single-stranded reverse ODN were mixed together and loaded onto the column. The absorbance at 280 nm showed elution of nucleic acid at two main peaks, one at 13.5 ml (peak 1) and the other at 15.5 ml (peak 2) (Figure 5.3). A sample of fractions corresponding to each peak were analysed by native PAGE. The first peak corresponded to the dsDNA and the second predominantly to ssDNA, but this peak also contained some dsDNA (Figure 5.3). Therefore, FPLC using the Superdex 200 column can separate a 40 bp duplex from 40 nt ssDNA.

5.2.1.3. Size exclusion FPLC can be used to separate R:D60 from ssRNA60 and ssDNA60

To establish if the largest hybrid formed by annealing oligonucleotides could be successfully fractionated from contaminating nucleic acids, R:D60 was run through the Superdex 200 column. Similar to the 40 bp dsDNA/ssDNA mixture, two main peaks were observed (Figure 5.4). Peaks 1 and 2 contain R:D60 and the contaminating band at approximately 100 bp. Peak 3, eluting between 13.0 – 13.5 ml exclusively contains R:D60, whilst peak 4 (15.0 – 15.5 ml) contains a mixture of R:D60 and ssDNA60 (Figure 5.4). Multiple repeats using different batches of annealed R:D60 consistently gave the two major peaks, always eluting at exactly the same volumes.

The fractions for each peak were concentrated by ethanol precipitation and resuspended in nuclease-free TE. On average, 55% of the input was recovered as R:D60 purified from contaminants.

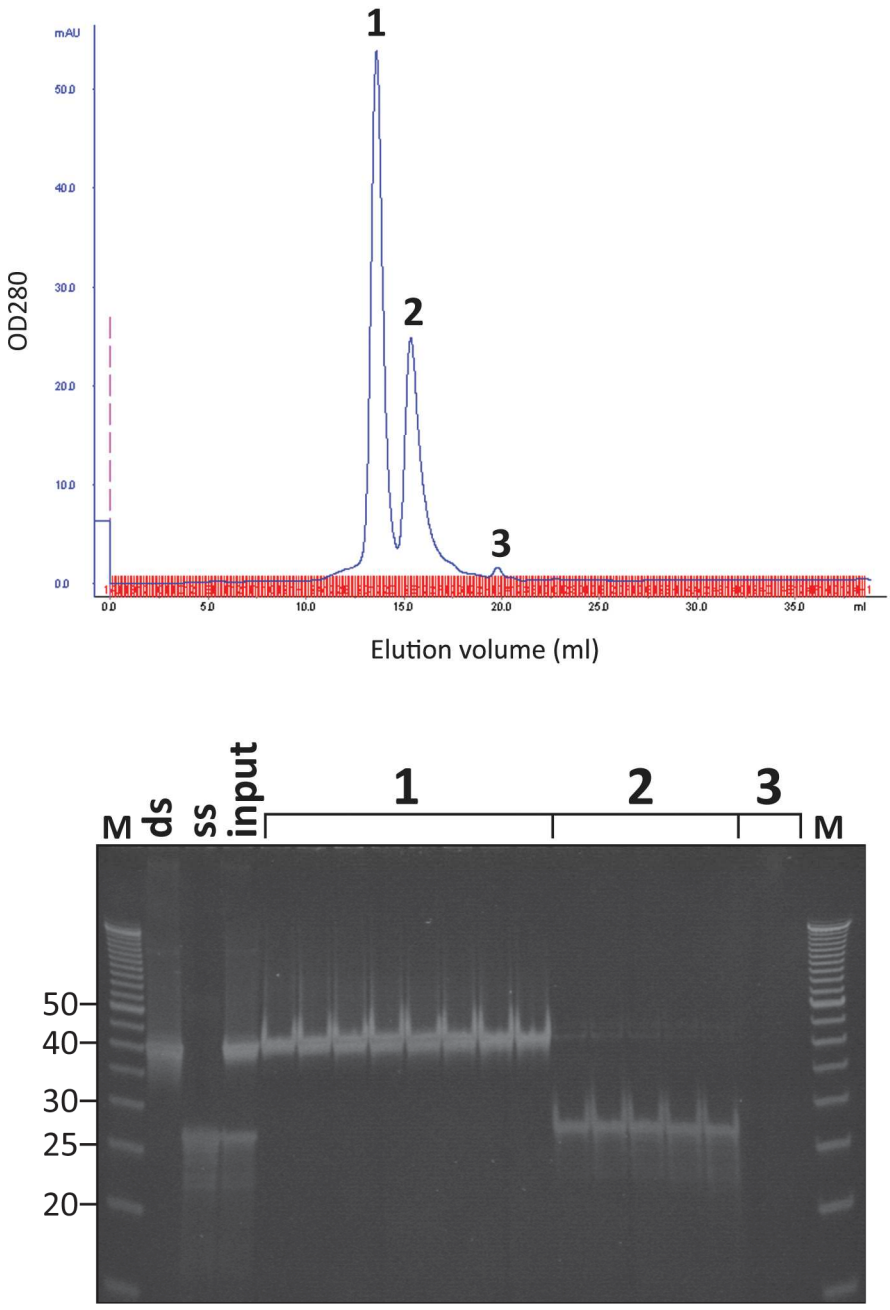


Figure 5.3. Separation of a 40 bp D-R-D:DNA duplex and 40 nt ssDNA.
100 µl of a mixture of equimolar amounts of nucleic acids (corresponding to 48.5 µg of double stranded and 24.3 µg of single-stranded) was injected into the column. Nucleic acid was eluted using nuclease-free buffer (60 mM KCl 50 mM Tris pH 8.0) at a flow rate of 0.4 ml/min. 100 µl fractions were collected and 10 µl of each fraction was electrophoresed on a 15% native polyacrylamide gel. Fractions loaded were from 13.2 – 14.0 ml (peak 1), 15.1 – 15.6 ml (peak 2) and 19.7 – 19.9 ml (peak 3). 200 ng each of dsDNA, ssDNA and the input was loaded as a control. 200 ng 5 bp ladder (M) was included as a size marker. The gel was stained with Sybr Gold and imaged using a UV transilluminator. Representative of duplicate FPLC runs using this duplex.

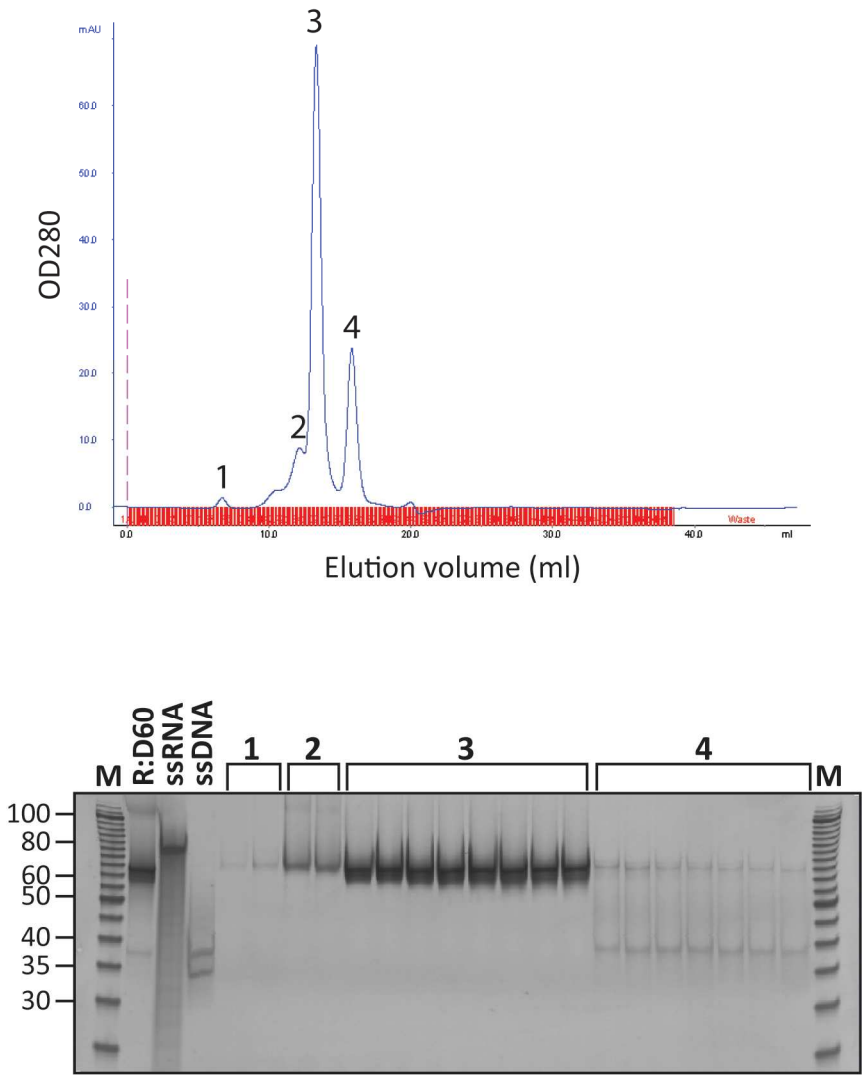


Figure 5.4. Gel filtration to separate R:D60 from contaminating nucleic acids.
72 µg of annealed R:D60 was fractionated by FPLC using the Superdex 200 column, eluting with nuclease-free buffer (60 mM KCl 50 mM Tris pH 8.0) at a flow rate of 0.4 ml/min. 100 µl fractions were collected and 10 µl of each fraction was analysed by native PAGE using a 15% gel. Fractions loaded were 10.3 – 10.5 ml (peak 1), 12.0 – 12.2 ml (peak 2), 12.9 – 13.7 ml (peak 3) and 15.4 – 16.2 ml (peak 4). 200 ng each of annealed R:D60, ssRNA60 and ssDNA60 were also loaded, and 200 ng 5bp ladder (M) was used as a size control. The gel was stained with Sybr Gold and imaged using a phosphorimager. Note the lower band in the ssDNA input lane is a by-product of incomplete purification of the oligonucleotide by the commercial suppliers. Representative of four independent runs.

5.2.1.4. Fractionation of 12, 18 and 20 bp RNA:DNA hybrids using FPLC

Size exclusion chromatography of the hybrids R:D12, R:D18 and R:D20 was performed as described for R:D60. As shown in Figure 5.5, each hybrid could be purified away from ssDNA contaminants. However, the purity of the hybrid fraction varied.

R:D18 (peak1, Figure 5.5 **b**) could be completely purified from ssDNA18 (peak 2, Figure 5.5 **b**). In addition, ssRNA18, which co-migrated at approximately 15 bp with the dsDNA size marker (see Chapter 4, Figure 4.2) is not visible in the fraction from peak 1. In contrast, the main R:D12-containing fraction (peak 1, Figure 5.5 **a**) also contained traces of ssDNA12 (band below 10 bp, Figure 5.5 **a**).

PAGE analysis of R:D20 before FPLC purification revealed the presence of multiple bands which migrated faster than the expected 20 bp RNA:DNA hybrid, visualised as a smear below the band at 20 bp (Figure 5.5 **c**, lane 3 'R:D20'). This could suggest that annealing of the single-stranded components had resulted in the generation of numerous additional shorter nucleic acid species. Alternatively, the repetitive sequence of this hybrid implies that these nucleic acids could have been RNA:DNA hybrid-containing partial duplexes. Such nucleic acid species would be generated if the two strands misaligned when they re-annealed, resulting in a duplex with single-stranded regions at either end. FPLC failed to separate these bands from the 20bp hybrid, probably due to their close proximity in size (Figure 5.5, peak 3),

To identify the composition of the nucleic acid species present in FPLC-purified R:D20, a sample of the fractions from peak 3 (Figure 5.5) were digested with RNase H1 (Figure 5.5 **d**). The majority of R:D20 was digested, including the smaller bands, leaving ssDNA20 which co-migrates with the dsDNA marker at approximately 15 bp. However, a band migrating at 20 bp remained. Incubation with greater concentrations of RNase H also failed to degrade this hybrid, indicating that is not simply undigested hybrid or hybrid-containing partial duplexes.

Further reactions using RNase I_f to digest ssRNA and Exo I to digest ssDNA confirmed that this band was ssRNA20 (Figure 5.5 **e**). As shown in Figure 5.5 **c** and **e**, ssRNA20 and R:D20 migrated at approximately the same size on a native polyacrylamide gel. It can be assumed that this overlap in migration resulted in a failure to separate ssRNA20 from R:D20 using size exclusion chromatography. Consequently, low levels of contaminating ssRNA20 were present in FPLC-purified R:D20. This indicated a limitation of the use of the Superdex 200 column for the separation of ssNA and dsNA of the same length. Size exclusion chromatography using a Superdex 75 column may be more appropriate for shorter hybrids (Kim *et al.*, 2007). The combination of the contamination of R:D20 with smaller RNA:DNA hybrid-containing nucleic acid species and ssRNA meant that FPLC-purified R:D20 was insufficiently pure for use in downstream experiments.

The migration of nucleic acids through the gel matrix of the Superdex 200 column, similar to their electrophoretic migration through native polyacrylamide gels, is dependent not only on length but also on the degree of secondary structure formed, particularly for ssNA. This in turn is influenced by the sequence of the nucleic acid, therefore the suitability of this technique has to be assessed for each hybrid, as demonstrated in Figure 5.5. The optimal separation of dsNA from ssNA using this column was obtained using longer duplexes; the 40 bp D-R-D:DNA duplex and R:D60 (Figures 5.3 and 5.4). Therefore, further characterisation of FPLC-purified R:D60 was performed to confirm it was a homogeneous RNA:DNA hybrid.

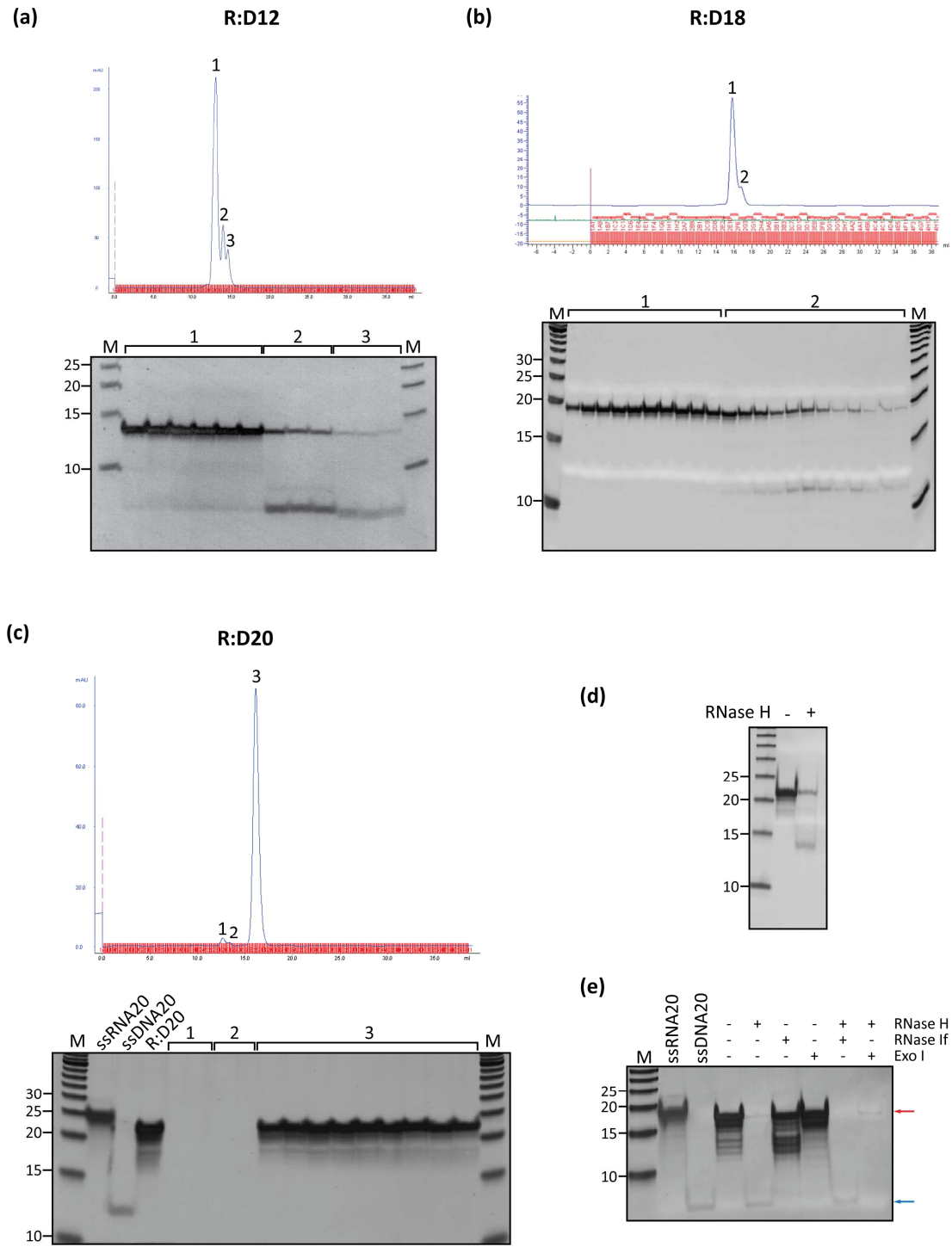


Figure 5.5. FPLC purification of R:D12, R:D18 and R:D20.
R:D12, R:D18 and R:D20 were FPLC purified using a Superdex 200 column as described for R:D60. Nucleic acids were electrophoresed on 20% native polyacrylamide gels, stained with Sybr Gold and imaged using a phosphorimager. 200ng 5bp ladder (M) was included as a size marker. **(a)** Fractions loaded were 12.6 – 13.2 ml (peak 1), 13.8 – 14.1 ml (peak 2) and 14.4 – 14.7 ml (peak 3). **(b)** Fractions loaded were 15.4 – 16.4 ml (peak 1) and 15.5 – 16.7 ml (peak 2). Note faint white lines at approximately 12 bp and 23 bp are the BPB and XC

reference dyes present in the sample loading buffer. **(c)** Fractions loaded were 12.6 – 12.8 ml (peak 1), 13.3 – 13.5 ml (peak 2) and 15.7 – 16.6 ml (peak 3). 250 ng each of ssRNA20, ssDNA20 and R:D20 input were electrophoresed alongside the fractions. **(d)** 500 ng (equivalent to 75 pmol) of FPLC-purified R:D20 was incubated in the presence or absence of 10 U *E. coli* RNase H in 20 µl reactions. Half of the reaction was electrophoresed on a 15% native polyacrylamide gel and stained with Sybr Gold. **(e)** 500 ng FPLC-purified R:D20 was incubated with 10 U RNase H, 10 U RNase I_f and 20 U Exo I as indicated. Half of each reaction was loaded onto a 15% native polyacrylamide gel, with 250 ng each of ssRNA20 and ssDNA20 included for reference. Band were visualised by Sybr Gold staining. ssRNA is indicated by a red arrow and ssDNA by a blue arrow (note bands are faint due to co-migration with the reference dyes in the loading buffer).

5.2.2. Validation of FPLC-purified R:D60

FPLC-purified R:D60 was digested with RNase H1 (Figure 5.6). As expected, the RNA strand was hydrolysed, resulting in a band which co-migrated with the dsDNA size marker at approximately 40 bp. This corresponded to the size of ssDNA60 (Figure 5.4). The nucleic acids were transferred to a positively charged nylon membrane which was probed with the RNA:DNA hybrid-specific S9.6 antibody. In the minus RNase H1 lane, signal corresponding to the RNA:DNA hybrid was detected which was absent in the sample treated with RNase H1 (Figure 5.6). This confirmed that the FPLC-purified 60 bp product generated by annealing complementary oligonucleotides is an RNA:DNA hybrid.

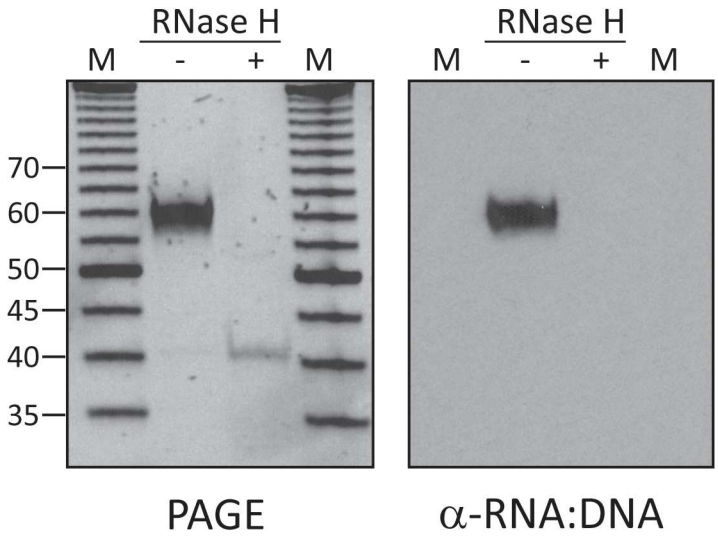


Figure 5.6. Characterisation of R:D60 using RNase H digestion and the S9.6 antibody. Digestion of R:D60 with RNase H resulted in a single band which migrated at the same size as ssDNA60. 200 ng FPLC-purified R:D60 was incubated in the presence or absence of 10 U *E. coli* RNase H1 for 1 h at 37°C. Half the volume of each reaction was electrophoresed on a 1 mm thick 12% native polyacrylamide gel, stained with Sybr Gold and imaged on a UV transilluminator at 254 nm. 300ng 5 bp dsDNA ladder was included as a size standard. Nucleic acids were then transferred to a positively-charged nylon membrane, which was probed with purified S9.6 antibody (1:1000) as described in Chapter 2.

5.2.3. Estimation of ssNA contamination of FPLC-purified R:D60

As an additional quality control step, each batch of R:D60 was electrophoresed alongside known amounts of ssNA. This allowed an estimation of any ssNA contamination. Most batches did not contain any detectable ssNA, however the hybrid shown in Figure 5.7 did contain some ssDNA60, most likely as a result of collecting fractions outside of the main peak. From the gel, it can be estimated that this batch contains approximately 6 ng of ssDNA60 (indicated by the blue arrow) per 100 ng of hybrid. Repeating the column purification could be used to eradicate this contamination.

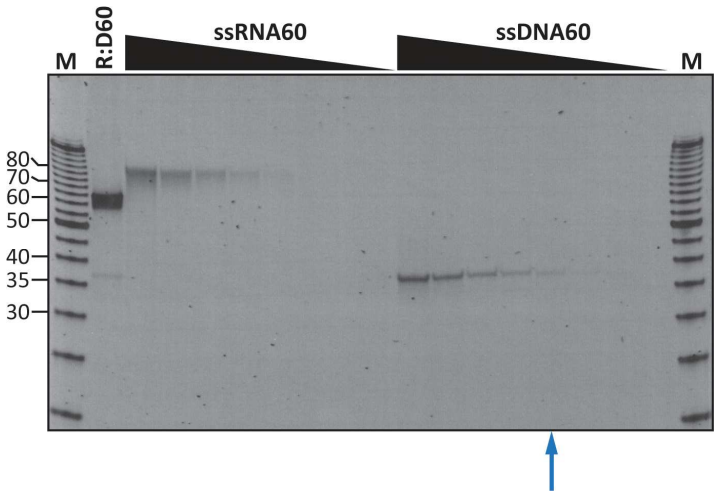


Figure 5.7. Quality controls step to quantify contamination of R:D60 by ssNA.
100 ng R:D60 was electrophoresed on a 15% native polyacrylamide gel with decreasing amounts of ssRNA60 and ssDNA60 (100, 50, 25, 12.5, 6.25, 3.1, 1.6, 0.8 ng). 200 ng 5bp ladder (M) was included as a size control. The gel was stained with Sybr Gold and imaged using a transilluminator. The blue arrow indicates the lane with 6.25 ng of ssDNA60.

5.3. The production and validation of a long RNA:DNA hybrid

The production of short (≤ 60 bp) RNA:DNA hybrids by the annealing of complementary oligoribonucleotides, as described in Section 5.2, provided a simple method to generate large amounts of RNA:DNA hybrid for subsequent testing in cells. However, it is possible that disease-causing endogenous RNA:DNA hybrids are much longer than those produced using this method. R loops can be up to a kb in size (Yu *et al.*, 2003). Additionally, hybrids formed from viral nucleic acids may be longer than 60 bp, for example the hybrids present at the origin of replication of human CMV (Prichard *et al.*, 1998). The source of endogenous RNA:DNA hybrids likely to accumulate in AGS patients has not been established, therefore it they are relatively long, a hybrid-detecting PRR may not be stimulated by the short hybrids described in Section 5.2 and consequently an immune response would not be triggered.

The length of RNA:DNA hybrids produced by annealing complementary strands is limited by the size of oligonucleotides that can be generated. Solid-support synthesis of ODNs between 100 nt and 140 nt long by commercial suppliers is possible but is severely limited by the coupling efficiency during the synthesis cycle. Additionally, the relative proportion of incomplete oligonucleotide by-products increases with the length of the oligonucleotide, so even at a very high coupling efficiency (99%), synthesis of an oligonucleotide > 100 nt generates a relatively low percentage of full-length product which requires extremely efficient purification from the unwanted by-products. The low yield of oligonucleotide recovered made this option impractical for this project.

Reverse transcription of the ssRNA viral genome is a key step in the lifecycle of retroviruses, generating an RNA:DNA hybrid intermediate. *In vitro*, ssRNA of a range of sizes can be synthesised using *in vitro* transcription (IVT), which could in theory be reverse transcribed to form an RNA:DNA hybrid. Therefore, to investigate the practicality of this method for generating hybrids suitable for testing in cells, a region of the CMV promoter sequence from the pEGFP-N1 plasmid (Clontech) was chosen as a template for IVT.

5.3.1. Strategy to generate an RNA:DNA hybrid using reverse transcription

The generation of long hybrids using reverse transcription involved a three step technique (Figure 5.8 **a**). Firstly, the chosen sequence was amplified using a high-fidelity DNA polymerase. The forward primer contained a minimal T7 polymerase promoter sequence (18 nt) followed by one extra guanosine to enhance transcription efficiency. The dsDNA PCR product was used as a template for IVT using T7 RNA polymerase. The resulting ssRNA was reverse transcribed with an RNase H-deficient reverse transcriptase (RT) using the reverse PCR primer. The nucleic acid products from each stage of this process for R:D188 are shown in Figure 5.8 **b**.

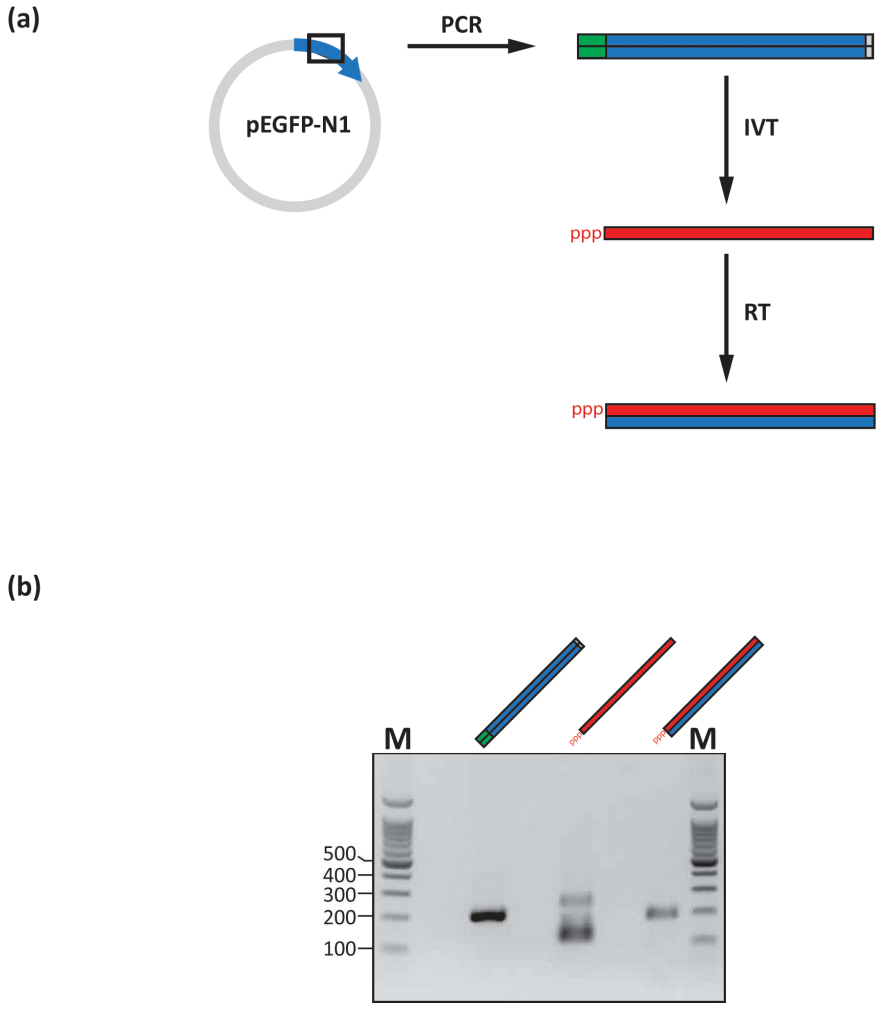


Figure 5.8. The generation of a 188 bp RNA:DNA hybrid by reverse transcription.
(a) Schematic diagram showing the process of generating R:D188. A region of the CMV promoter (blue) encoded within an endotoxin-free maxiprep of pEGFP-N1 was amplified using PCR. The forward primer contained a T7 promoter sequence (green) and the reverse contained an *Eco*RI site (grey). The PCR product was gel extracted and used as a template for *in vitro* transcription (IVT) using T7 RNA polymerase. ssRNA was DNase I treated to remove all traces of dsDNA template and PAGE purified to remove aberrant transcription products. RNA was then used as a template for reverse transcription (RT) using a sequence specific reverse primer. The resulting hybrid was 188 bp long and contained a 5′-ppp. **(b)** dsDNA, ssRNA and RNA:DNA hybrid were electrophoresed on a 1.5% agarose gel containing ethidium bromide.

5.3.2. Optimisation of method to generate ssRNA

The method described above was refined to reduce potential contamination of the hybrid with other nucleic acids, particularly 5′-ppp RNA which is potentially a RIG-I/TLR7 ligand.

The MegaShortScript T7 polymerase kit (Ambion) was successfully used to produce high yields of 188 nt long IVT RNA, using 200 ng of 206 bp (1.47 pmol) dsDNA template, an average of 135 μg of 188 nt (2.1 nmol) ssRNA was produced. Electrophoresis of the RNA on non-denaturing gels revealed the presence of more than one band (Figure 5.9 **a**). Mfold analysis (Zuker, 2003) of the RNA sequence indicated that it can form extensive secondary structures (Figure 5.9 **a**). However, subsequent analysis of the IVT RNA by denaturing PAGE suggested that the additional bands were multiple RNA species rather than secondary structure.

The generation of large amounts of aberrant RNA products during IVT with T7 RNA polymerase has been described previously (Cazenave and Uhlenbeck, 1994; Schenborn and Mierendorf, 1985; Schlee *et al.*, 2009b; Schmidt *et al.*, 2009; Triana-Alonso *et al.*, 1995). The enzyme is able to extend the 3′ end of the newly transcribed RNA transcript using RNA as a primer. This can occur during foldback of the nascent transcript or from the annealing of another RNA molecule. Both result in the formation of a complementary double-stranded template region and the production of longer RNA transcripts, with varying degrees of extension (Cazenave and Uhlenbeck, 1994; Triana-Alonso *et al.*, 1995). Failure to purify IVT RNA from 3′ extended by-products resulted in confusion over the true ligand for the RNA helicase RIG-I (Schlee *et al.*, 2009b; Schmidt *et al.*, 2009). The use of such RNA as a template for reverse transcription could generate multiple RNA:DNA hybrid species, which may not be fully duplex, as a sequence specific reverse primer is used during the reverse transcription step. Therefore, the IVT reaction was modified to limit the possibility of this.

The production of these aberrant RNA transcripts can be inhibited by reducing the reaction time, reducing the concentration of T7 polymerase used and decreasing the

UTP concentration (Triana-Alonso *et al.*, 1995). Additionally, the use of a dsDNA template with a 5′ overhang has also been shown to limit the formation of unwanted RNA products (Schenborn and Mierendorf, 1985). Therefore, an *Eco*RI restriction enzyme site was incorporated into the reverse DNA primer used for PCR amplification of the template (Figure 5.8 **a**) and digestion performed prior to the IVT reaction. IVT RNA was subsequently purified using denaturing PAGE to further ensure the correct transcript was used (Figure 5.9 **b**).

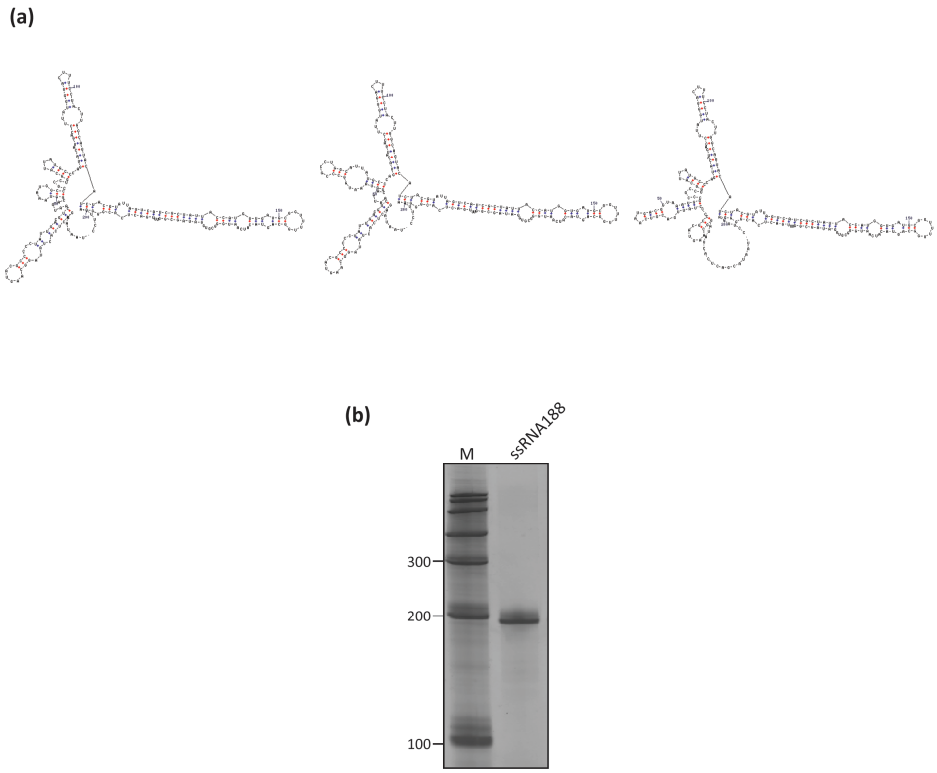


Figure 5.9. Analysis of IVT ssRNA188 derived from a CMV promoter sequence.
(a) IVT ssRNA188 was predicted to contain extensive regions of secondary structure in non-denaturing conditions. Three different higher order structures predicted by Mfold (Zuker, 2003) are shown. **(b)** 250 ng of PAGE purified ssRNA188 was electrophoresed on an 8% denaturing polyacrylamide gel, which was stained with Sybr Gold and imaged using a phosphorimager. 2 µl of of RiboRuler Low Range RNA ladder was included as a size marker (M), numbers indicate size in nt.

5.3.3. Validation of R:D188

The 188 bp RNA:DNA hybrid generated by reverse transcription was analysed to validate its nature and confirm the purity of the hybrid. Digestion with RNase H2 resulted in a down-shifted band, which appeared red with acridine orange staining which confirmed that it was single-stranded (Figure 5.10 **a**).

R:D188 was detected by the S9.6 antibody but RNase H2-digested hybrid was not (Figure 5.10 **b**). Probing of the same samples with a commercial monoclonal antibody against ssDNA, revealed that RNase H2-digested R:D188 contains ssDNA, which confirmed that the RNA strand of the hybrid had been hydrolysed (Figure 5.10 **b**). However, there was also signal in the undigested R:D188, which suggested the presence of ssNA. Notably, the antibody did not detect a control RNA:DNA hybrid thereby excluding the possibility of non-specific binding (negative control, α-ssDNA blot, Figure 5.10 **b**). However, the antibody also detected signal in the minus RT control, which does not contain any ssDNA, other than the 25 nt ssDNA primer used in the reverse transcription reaction, which would be used up during the generation of the complementary DNA strand. It did however contain ssRNA: the template for reverse transcription.

Consequently, the specificity of the antibody for ssNA was characterised using 18 base oligonucleotides. In addition to the ssDNA oligonucleotide, the antibody also detected ssDNA containing a single ribonucleotide (ssD-R-D) and a ssRNA oligonucleotide (Figure 5.10 **c**). This suggested that the ssNA detected in R:D188 could have been IVT ssRNA template or ssDNA primers carried over from the reverse transcription reaction. Alternatively, the reverse transcriptase enzyme may dissociate from the primer-template, generating incomplete hybrids and ssNA; it has been shown that the presence of certain sequences or secondary structures in the template causes the dissociation of HIV-1 reverse transcriptase *in vitro* (Abbotts *et al.*, 1993; DeStefano *et al.*, 1992; Klarmann *et al.*, 1993). If the enzyme does not reassociate with the RNA template, transcription is aborted, which may generate incomplete products.

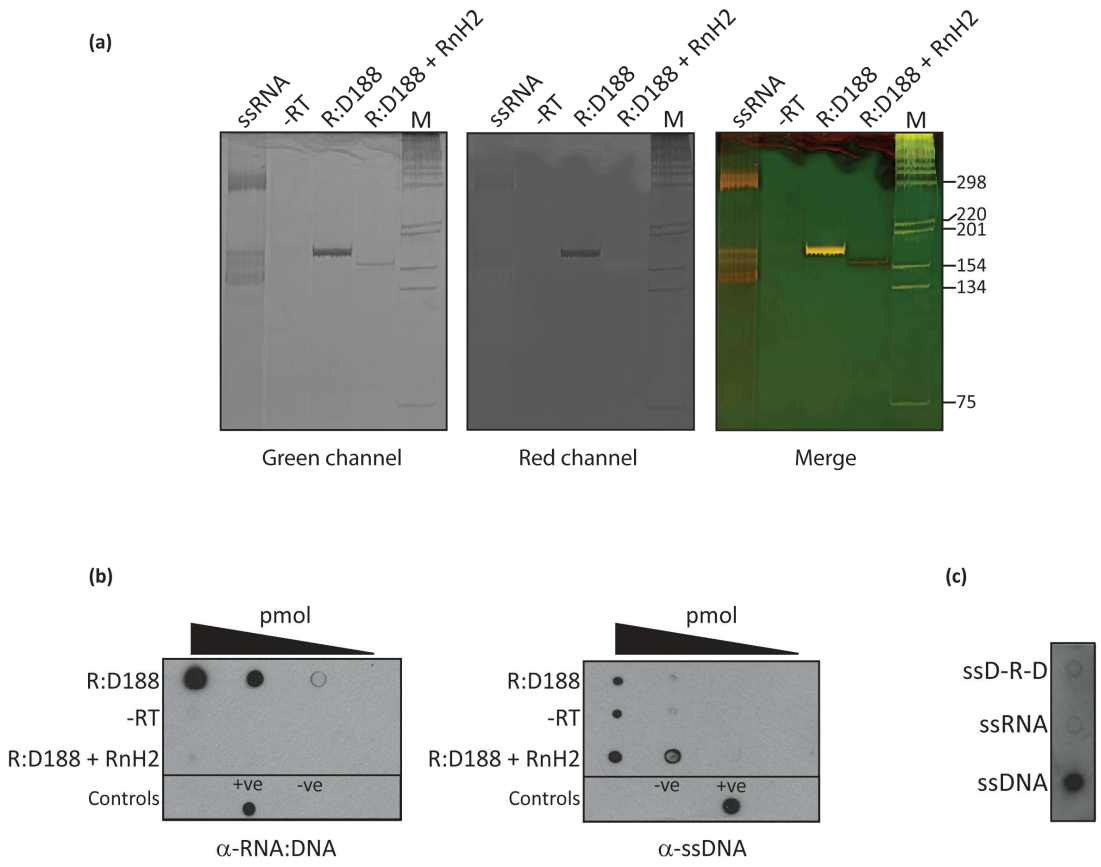


Figure 5.10. Validation of R:D188.
(a) 10 µg R:D188 was incubated in the presence or absence of 770 ng recombinant human RNase H2 (from Rachel White, MRC HGU) for 1 h at 37°C. 1 µg or equivalent volumes of each of unpurified IVT ssRNA, a minus reverse transcriptase (-RT) control, undigested R:D188 and RNase H2-digested R:D188 were electrophoresed on an 8% native polyacrylamide gel. 500 ng of 1 kb ladder was included as a size control (M). The gel was stained with acridine orange and imaged using a phosphorimager. Images were taken at two different wavelengths (excitation at 532 nm with SHG laser (green channel) and excitation at 473 nm with He-Ne laser (red channel)). The two channels were digitally merged using IPLab software. Single-stranded nucleic acids stained red, and double-stranded green. **(b)** Validation of R:D188 using the S9.6 antibody. Decreasing amounts of R:D188 (16, 1.6, 0.16 and 0.016 pmol) and equivalent volumes of –RT and RNase H2-digested R:D188 were UV crosslinked to a nitrocellulose membrane which was probed with the S9.6 anti-RNA:DNA hybrid antibody. A duplicate membrane was probed with a commercial antibody against ssDNA. Controls for the antibodies were 20pmol R:D18 and 20pmol ssDNA18. **(c)** The ssDNA antibody detected 18nt oligonucleotides. Blot is representative of two independent experiments to verify the specificity of the anti-ssDNA antibody.

An additional complication of this hybrid-generating procedure is the DNA-dependent DNA polymerase activity in RT enzymes (Roth *et al.*, 1985). They are able to utilise both dsDNA and RNA:DNA hybrids as templates for transcription; therefore transcription of ssDNA from the RNA:DNA hybrid product could also have occurred.

The detailed analysis of R:D188 raised potential issues associated with the use of reverse transcription to generate long RNA:DNA hybrids. The stimulation of cells with nuclease-treated R:D188 (detailed in Chapter 6) also indicated potential contamination of the hybrid, further questioning the purity of hybrids generated using this methodology.

5.4. Alternative approaches to generating long RNA:DNA hybrids

The short RNA:DNA hybrids produced in Section 5.2 may fail to stimulate innate immunity due to their size, depending on which PRRs, if any, are capable of detecting hybrids. Additionally, long hybrids may be more physiologically relevant to AGS and consequently more likely to be immunogenic. The concerns over the use of reverse transcription to generate suitable RNA:DNA hybrids prompted the examination of methods to produce the two complementary strands independently and anneal them together, as described for hybrids of $\leq$ 60 bp. Once established, this methodology could be used to generate hybrids ranging in length from less than 100 bp to in excess of 500 bp, depending on the efficiencies of annealing longer nucleic acid. It could also be used to generate long hybrids of the same size with different sequences to determine if the detection of RNA:DNA hybrids by sensors of the innate immune system is both length- and sequence-dependent

As described in Section 5.3.2, large quantities of a homogeneous population of long ssRNA can be produced by gel purification of IVT reactions. Commercial production of ODNs longer than 100 nt is inefficient. Therefore, techniques to generate large amounts of long (188 nt) ssDNA were established and tested for suitability. An overview of each of the methods is summarised in this section.

5.4.1. The use of PCR to synthesise ssDNA

It is possible to preferentially amplify one strand of template DNA using an excess of one primer. This technique, termed asymmetric PCR, was originally described to produce an excess of ssDNA for sequencing or use as a hybridisation probe by using a 10 to 1000-fold molar excess of one primer over the other (Gyllensten and Erlich, 1988). In the early cycles of the PCR, the entire amplified product is dsDNA. After the limiting primer has been exhausted, an excess of ssDNA is produced with every cycle. This increases in a linear fashion, as opposed to the exponential production of dsDNA by regular PCR, and so a greater number of cycles are required to obtain sufficient ssDNA product (Gyllensten and Erlich, 1988). The products consist of dsDNA and an excess molar excess of ssDNA (Figure 5.11 **c**).

Primers were designed to amplify a 362 bp region of the CMV promoter (forward primer; CMVprom_F, reverse primer; CMVprom362_R) (Figure 5.11 **a** and **b**). The PCR product was purified by gel extraction and used as a template for asymmetric PCR, with an excess of the reverse primer (CMVprom188_R) (Fig. 5.11 **c**). A range of primer ratios, amount of template and number of amplification cycles were tested. As shown in Figure 5.11 **d**, the band at 188 bp increased with intensity as the number of cycles increased, however no additional band was resolved, which suggested that either ssDNA was not generated or that it co-migrated with dsDNA188, which would make gel purification of specifically the ssDNA difficult. Therefore, the use of asymmetric PCR to generate ssDNA proved unsuccessful using a range of experimental conditions.

Single-stranded DNA templates for cycle sequencing are frequently prepared by linear amplification, using only a single primer (Kaltenboeck *et al.*, 1992). Therefore, reactions were performed using the same 362 bp dsDNA template, completely omitting the forward primer, so the only product should be ssDNA from linear amplification of the reverse strand. Optimisation of reaction conditions revealed that

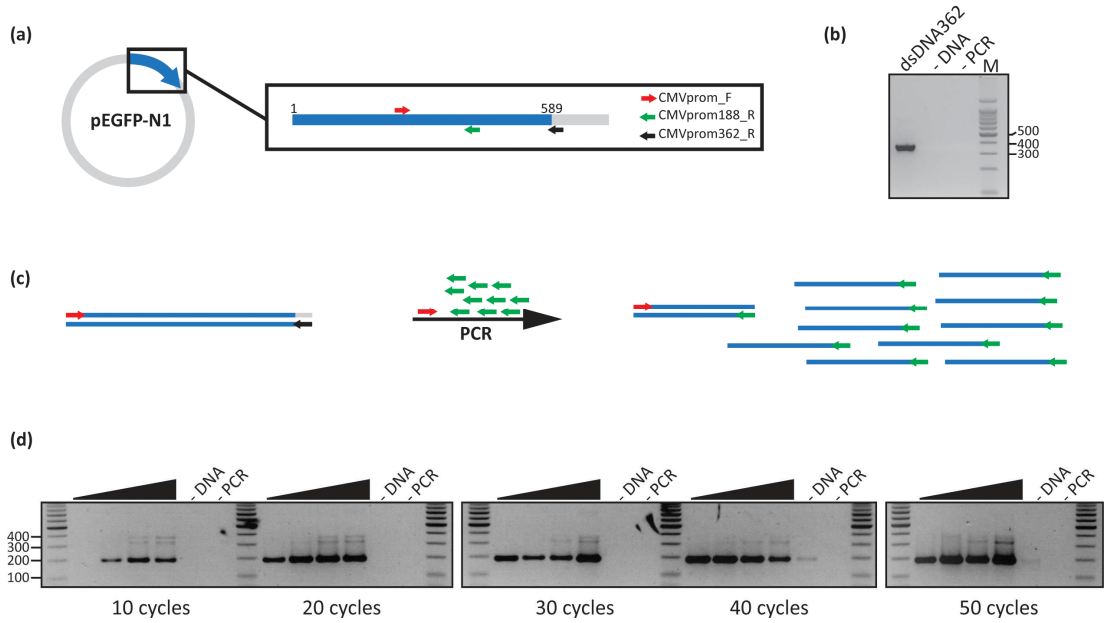


Figure 5.11. Asymmetric PCR to generate reverse strand ssDNA.
(a) Schematic showing location of primers within the CMV promoter region (blue) of pEGFP-N1. **(b)** A 362 bp region was amplified using the primers CMVprom_F (red) and CMVprom363_R (black). The PCR product (dsDNA362) was electrophoresed on a 1.5% agarose gel containing ethidium bromide alongside control PCR reactions (minus template (-DNA) and template plus water only (-PCR)). 200 ng of 100 bp ladder (M) was included as a size control. **(c)** Schematic of the principle of asymmetric PCR using a ten-fold excess of the reverse primer CMVprom188_R (green). **(d)** Example of the products generated using asymmetric PCR. Increasing amounts of dsDNA362 template were used (0.1, 1, 5, 10 ng), as indicated above each gel. The numbers of cycles were varied as indicated underneath each gel. Controls for PCR were included with each reaction (-DNA, -PCR). Equal volumes of PCR products were electrophoresed on a 1.5% agarose gel containing ethidium bromide, with 200 ng 100 bp ladder included as a size control.

a large number of cycles together with a low annealing temperature were required to generate sufficient product (indicated by blue asterisk, Figure 5.12 **a**).

A band migrating at approximately 150 bp was consistently produced (Figure 5.12 **b**). As ssDNA188 would be expected to migrate faster than dsDNA of the same size, this band was extracted from an agarose gel. Native PAGE analysis confirmed the successful purification of the correct band (Figure 5.12 **c**). However, heat denaturation of this product revealed that it consisted of two species (Figure 5.12 **c**), which suggested the presence of dsDNA. The failure of any products of the PCR to be digested by the ssDNA-specific nuclease Exo I confirmed that it was not ssDNA

(Figure 5.12 **d**). Therefore, under the conditions described, both asymmetric PCR and linear amplification using one primer were not suitable techniques for the aims of this method.

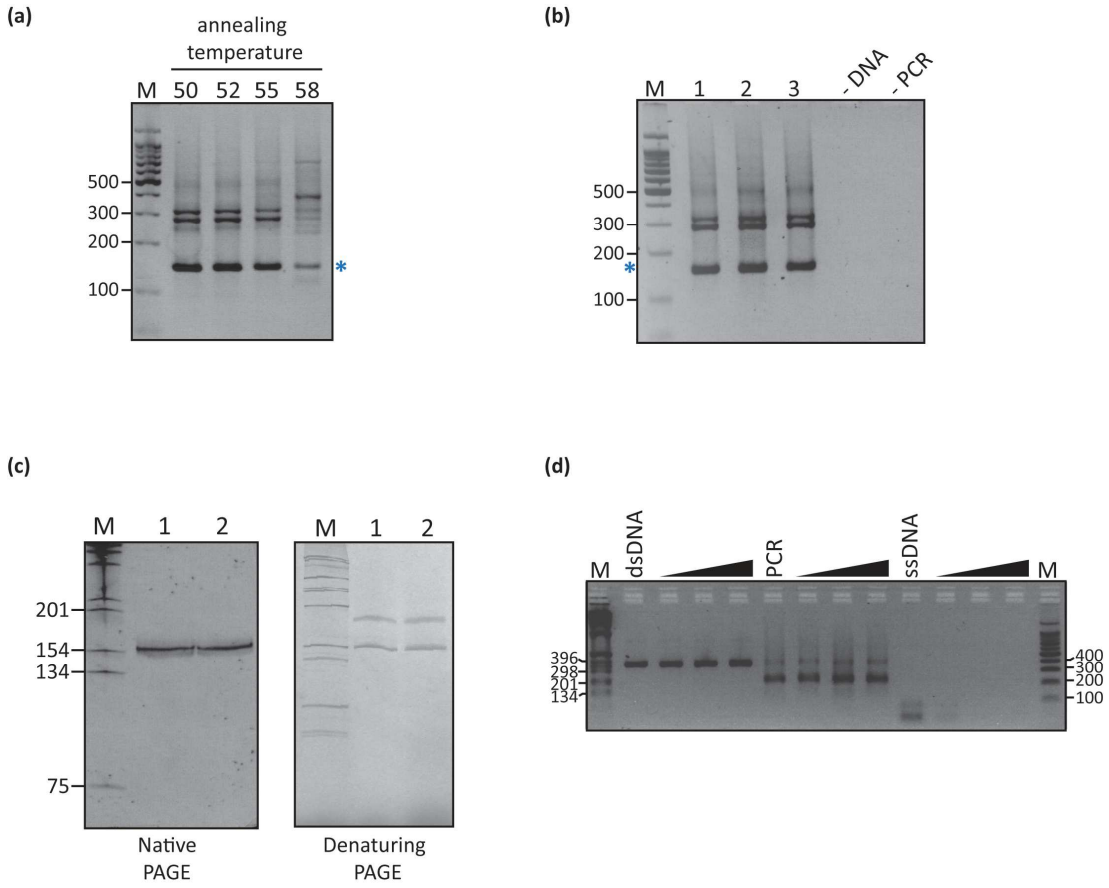


Figure 5.12. Reverse primer only PCR.
(a) Temperature gradient PCR. Products were electrophoresed on a 2% agarose gel containing ethidium bromide. The suspected ssDNA188 band is indicated by a blue asterisk. 200ng 100bp ladder (M) was included as a size marker. **(b)** PCR products generated from triplicate (1, 2 and 3) PCR reactions using an annealing temperature of 52°C and 50 cycles. PCR products were electrophoresed on a 2% agarose gel containing ethidium bromide. **(c)** PAGE analysis of the gel extracted band. Duplicate extracted bands were electrophoresed on 8% native PAGE (left) and 8% denaturing PAGE (right) gels. On both gels, 300 ng of 1 kb ladder (M) was included as a size marker. Gels were stained with the ssNA and dsNA stain Vistra Green and imaged using a phosphorimager. **(d)** Exo I digest of a PCR products. 200 ng each of dsDNA362 (Figure 5.11 **b**), reverse primer PCR (this figure, **(a)**) and a 65 nt ssDNA oligonucleotide were digested with increasing amounts of Exo I (10, 20, 50 U) for 20 min at 37°C. Products were electrophoresed on a 2% agarose gel containing ethidium bromide. The gels shown in **(d)** is representative of duplicate experiments.

5.4.2. Purification of a non-labelled strand of a DNA duplex

The second approach to the production of long ssDNA involved the biotin labelling of the forward strand by PCR (Figure 5.13 **a**). The 188 bp biotinylated PCR product was gel purified and bound to streptavidin. Techniques for separating the two strands to allow purification of the non-biotinylated reverse strand were then tested.

Firstly, the dsDNA-streptavidin complex was heated to 95^{o}C to denature the PCR product, without destroying the streptavidin-biotin bonds, before being subjected to electrophoresis using native PAGE. The additional size of the streptavidin molecule should cause the forward strand to supershift, allowing the unlabelled strand to be excised from the gel (Figure 5.13 **b**). The purification of ssDNA using this technique has been reported previously (Pagratis, 1996), however in this case it was found to be a very inefficient process, resulting in a very poor yield of ssDNA (Figure 5.13 **c**).

Alternative methods for purifying the non-biotinylated strand utilised immobilisation of the biotinylated PCR production onto streptavidin magnetic beads (Dynabeads, Invitrogen). Initially, the beads were heated to 95^{o}C and immediately removed from solution using a magnet. The reverse strand was then collected from the supernatant (Figure 5.14 **b**). Upon incubation with streptavidin beads, most of the biotinylated dsDNA successfully bound to the beads, as a relatively small proportion was detected in the non-bound fraction (NBF) (Figure 5.14 **c**). However, after heat denaturation of the duplex, some dsDNA was recovered from the supernatant (Figure 5.14 **c**). Additionally, the supernatant contained very low amounts of non-biotinylated ssDNA (indicated by a blue asterisk, Figure 5.14 **c**).

As heating to 95^{o}C in a non-deionising buffer should not break the biotin-streptavidin bond, this is likely to be unbiotinylated PCR product (despite the use of a HPLC-purified biotinylated forward primer for the initial PCR, Figure 5.14 **a**). Recovery of nucleic acids from the beads by heating to 95^{o}C in formamide revealed the presence of only dsDNA (Figure 5.14 **c**). Therefore, the production of large amounts of ssDNA using this method would be highly inefficient.

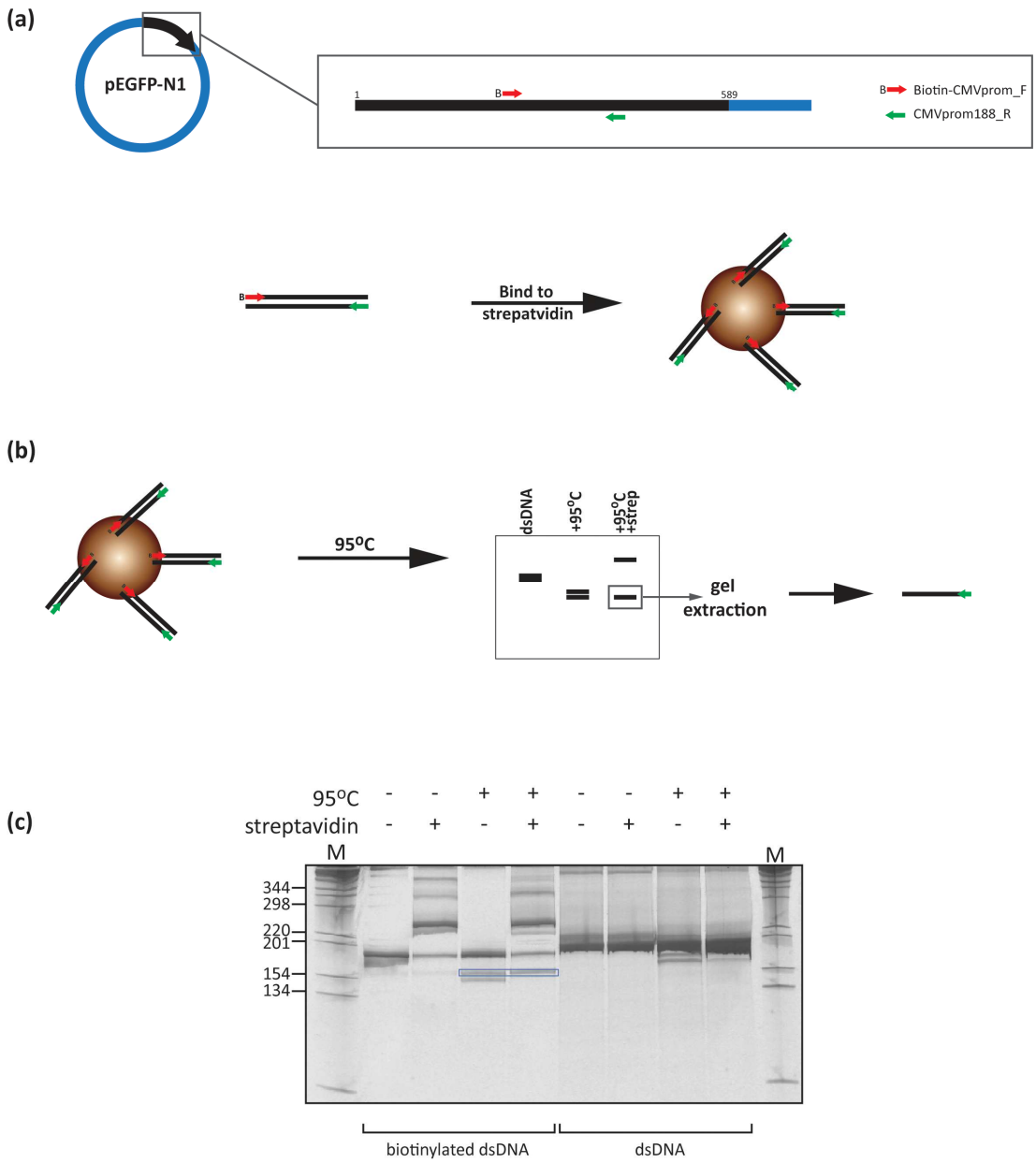


Figure 5.13. Streptavidin supershift to purify ssDNA.
Schematic representations of **(a)** amplification of a 188 bp region of the CMV promoter using a biotinylated forward primer and **(b)** purification of the non-biotinylated strand. Each streptavidin molecule has four biotin-binding sites (Chaiet and Wolf, 1964), although binding capacity can also be affected by the fragment length and salt concentration. **(c)** Biotinylated 188 bp PCR products and non-biotinylated controls (dsDNA) were incubated in the presence or absence of streptavidin-HRP for 30 min at room temperature, and then half were heated to 95°C for 5 min to denature the DNA before electrophoresis on an 8% native polyacrylamide gel. The gel was stained with Vistra Green and imaged using a phosphorimager. The non-biotinylated ssDNA strand is indicated with a blue box.

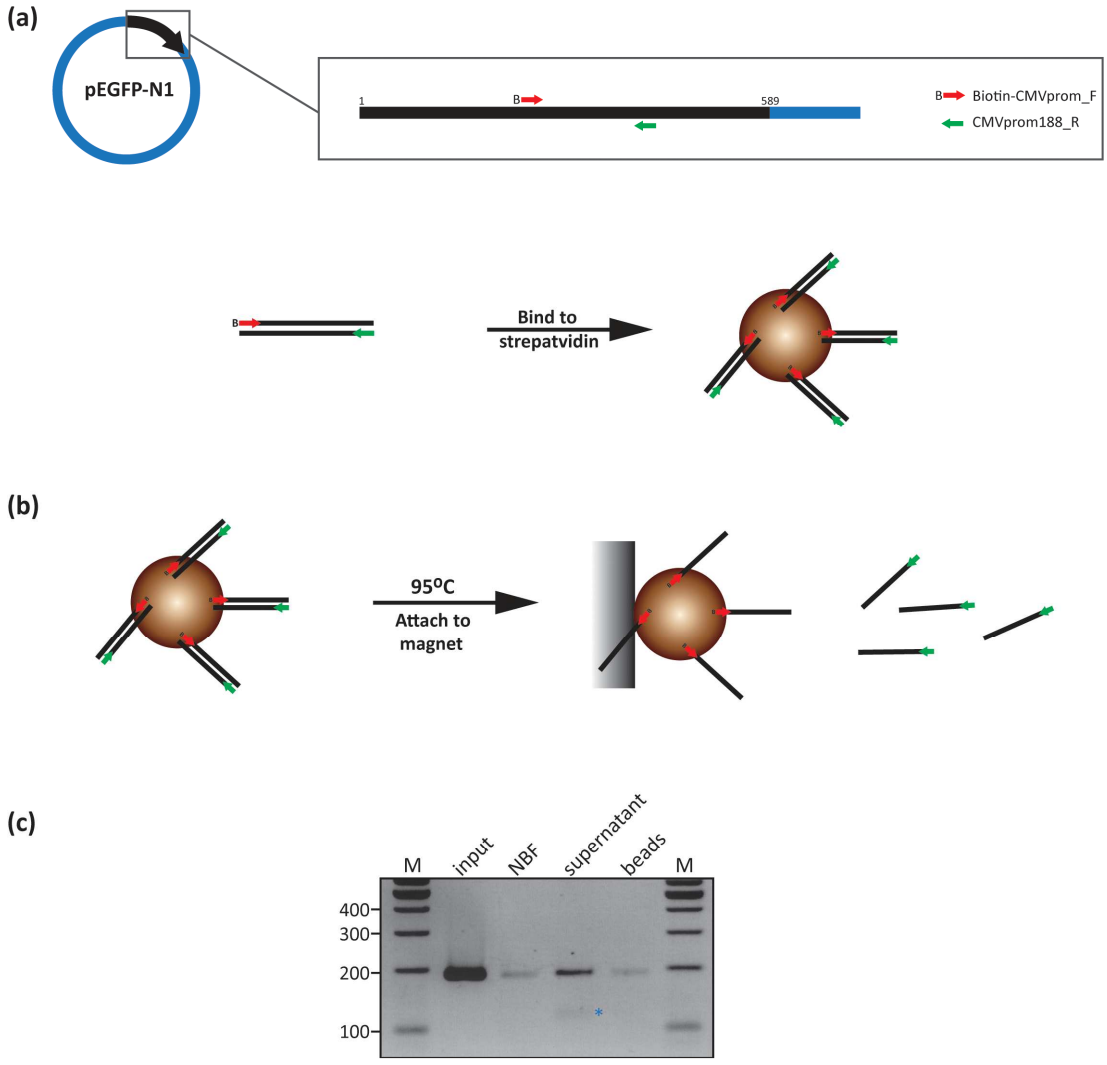


Figure 5.14. Heat denturation of streptavidin-bound dsDNA
(a) As for Figure 5.13. **(b)** dsDNA, bound to streptavidin Dynabeads via the biotinylated forward strand only, is heat denatured before the beads are immobilised onto a magnet. The non-biotinylated reverse strand is no longer attached to the bead and so can be collected from the supernatant. **(c)** 10% of the input dsDNA, and 15% each of the ethanol precipitated non-bound fraction (NBF), supernatant collected following heat denaturation and remaining DNA recovered from the beads (by heating to 90°C for 2 min in 95% formamide, 10 mM EDTA) were electrophoresed on a 2% agarose gel containing ethidium bromide. The faint band corresponding to the reverse ssDNA in the denatured fraction is highlighted with an asterisk.

Immobilised dsDNA was incubated with 100 mM NaOH and 200 mM NaCl (Figure 5.15 **a**). The use of NaOH and high NaCl concentrations to separate an 80 nt non-biotinylated strand from a biotinylated product complexed to streptavidin Dynabeads has been reported previously (Sidorov *et al.*, 2004). Nucleic acid was successfully recovered, neutralised and ethanol precipitated using this method (Figure 5.15 **b**). However, when electrophoresed, the eluted nucleic acid comprised of one band at the assumed size for ssDNA188 (migrating below the 154 bp band of the dsDNA size marker) and a number of higher molecular weight bands.

To identify the presence of any biotinylated nucleic acids in the recovered non-labelled ssDNA, the fractions were incubated with streptavidin to induce an electrophoretic shift. As shown in Figure 5.15 **b**, the presumed ssDNA band did not shift (lane 6), confirming it as non-biotinylated ssDNA. The same size product recovered from the beads did shift, confirming it was labelled with biotin (lane 8). The supernatant fraction (lane 5) also contained a band which migrated at the same size as the 188 bp dsDNA input (lane 1). Following incubation with streptavidin, this band also shifted to approximately 220 bp (lane 6), and so indicated the presence of biotinylated-dsDNA in the supernatant fraction.

This experiment provided further evidence of the presence of non-biotinylated dsDNA188 in the input DNA, as a proportion of it did not bind to the beads and so was recovered in the NBF (lane 3). This dsDNA species did not shift upon incubation with streptavidin (lane 4), therefore is not biotinylated.

As described, the nucleic acid fraction recovered in the supernatant contained a number of higher molecular weight bands (lanes 5 and 6, Figure 5.15 **b**), which could have been due to the formation of extensive secondary structure upon renaturation. To further identify these bands, all samples were electrophoresed on a denaturing PAGE gel (Figure 5.15 **c**). The higher molecular weight bands were not present under denaturing conditions, suggesting that they were a result of a different ssDNA conformational folding. However, whilst there was only one band in the lanes

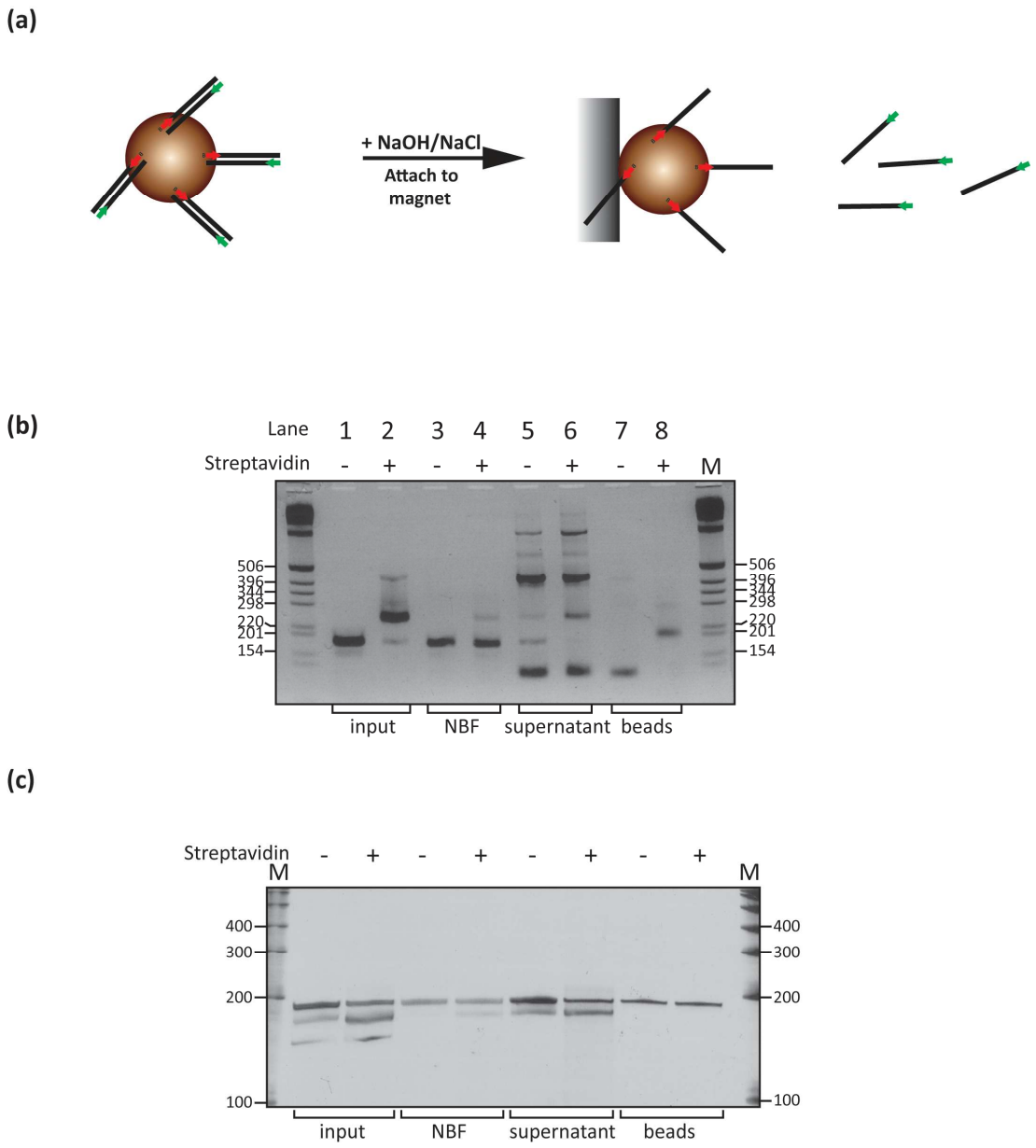


Figure 5.15. Sodium hydroxide denaturation of streptavidin-bound dsDNA.
Biotinylated dsDNA was generated and immobilised onto streptavidin as shown in Figure 5.15. **(a)** Dynabeads are incubated for 3 x 10 min in 100 mM NaOH/200 mM NaCl to denatured the dsDNA and allows recovery of the non-biotinylated strand. **(b)** The input, non-bound fraction (NBF), denatured fraction and biotinylated fraction eluted from the beads were incubated in the presence of absence of streptavidin-HRP. 5% of input and 15% of the fractions were electrophoresed on a 2% agarose gel containing ethidium bromide. 200 ng of 1 kb ladder was included (M). **(c)** The same volume of samples loaded in **(b)** were electrophoresed on a 10% denaturing polyacrylamide gel. 2 µl of the low range RiboRuler RNA ladder was included as a size guide (labelled sizes are in nt). The gel was stained with Sybr Gold and imaged using a transilluminator. Representative of duplicate experiments.

containing ssDNA from the beads, there were two bands present in the ssDNA recovered from the supernatant (Figure 5.15 **c**). Therefore, ssDNA recovered from supernatant following denaturation with NaOH/NaCl was not a homogeneous species, and so would require additional purification before being annealed to a complementary RNA strand.

5.4.3. Digestion of the forward strand by λ exonuclease

Finally, as an alternative approach to the isolation of reverse strand ssDNA, the same 188 bp region of the CMV promoter was amplified using a 5′-end phosphorylated forward primer (Figure 5.16 **a**). The PCR product was gel extracted and incubated with λ exonuclease, which catalyses the removal of 5′ mononucleotides from 5′-phosphorylated dsDNA in a 5′ → 3′ direction. Therefore, when incubated with a DNA duplex containing a 5′-phosphorylated strand, the enzyme selectively digests only this strand, leaving the non-phosphorylated ssDNA strand intact (Kujau and Wolfl, 1997).

The concentration of λ exonuclease used in each reaction required optimisation, as the enzyme was found to cleave the non-phosphorylated strand when used at high concentrations (Figure 5.16 **b**). However, when used at lower concentrations, a greater proportion of dsDNA remained undigested (Figure 5.16 **b**). The product which migrated at approximately 150 bp was completely degraded by Exo I, which confirmed it was ssDNA (Figure 5.16 **c**). Despite extensive optimisation of incubation times, and concentrations of both the template and enzyme, a point where the entire phosphorylated strand is digested without degradation of the non-phosphorylated strand could not be established.

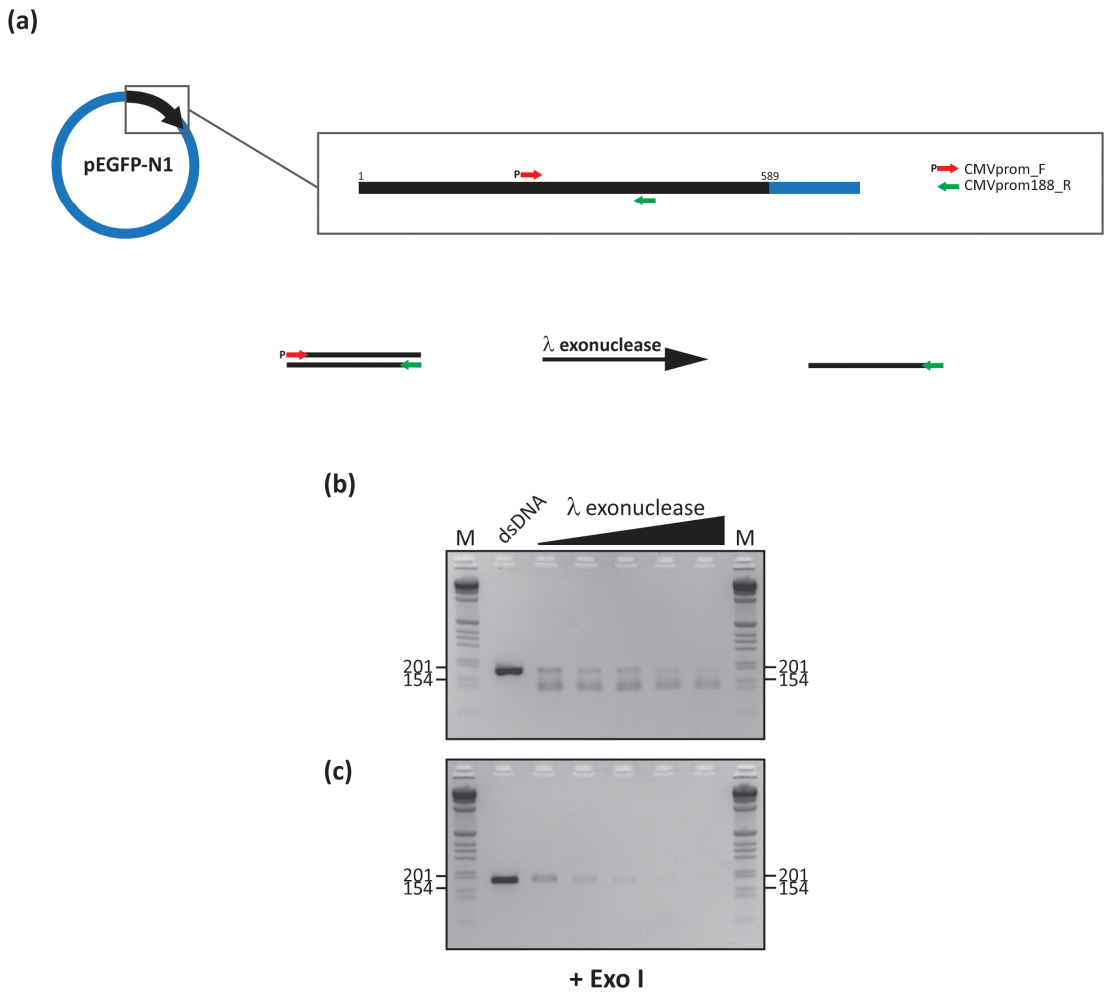


Figure 5.16. The use of λ exonuclease to digest one strand of a DNA duplex.
(a) Schematic representation of the amplification of a 188 bp region of the CMV promoter using a 5′-phosphorylated forward primer. λ exonuclease will digest the phosphorylated strand, leaving the reverse strand intact. **(b)** 1 µg of gel extracted PCR products were incubated with increasing amounts of λ exonuclease (0, 1, 2.5, 5, 7.5, 10 U) in 10 µl (therefore final DNA concentration was 100 ng/µl). 2 µl of each reaction was electrophoresed on a 2% agarose gel containing ethidium bromide. (c) 4 µl of each reaction from (b) was digested with 20 U of Exo I and electrophoresed on a 2% agarose gel containing ethidium bromide. 200 ng of 1 kb ladder was included as a size marker (M). Experiment is representative of multiple replicates to optimise enzyme and template concentrations.

5.4.4. Summary of methods for synthesising 188 nt ssDNA

The generation of ssDNA by asymmetric and reverse primer PCR was largely unsuccessful, as it was not possible to identify the ssDNA PCR products with confidence. The amplification of 188 bp dsDNA using modified forward strand primers yielded more encouraging results. However, the biotin-streptavidin electrophoretic shift would not be an entirely practical method, as it would be likely that only very low yields of ssDNA would be recovered. Additionally, the contamination of reverse strand ssDNA denatured from streptavidin-bound dsDNA with both biotinylated and non-biotinylated PCR products is a further disadvantage of this technique. The use of low concentrations of λ exonuclease to degrade the forward strand of dsDNA was found to be the most appropriate method, and would be a practical technique to scale up for the generation of large quantities of product. However, the ssDNA required purification to separate it from any undigested dsDNA, therefore techniques to purify 188 nt ssDNA from 188 bp dsDNA were assessed.

5.5. The purification of long ssDNA

The purification of DNA from agarose or polyacrylamide gels is commonly used in many protocols. Excision and purification of the ssDNA product of λ exonuclease digestion from both agarose and native polyacrylamide gels was tested. Whilst pure ssDNA could be obtained following elution from polyacrylamide gels, the yields were low, therefore the generation of enough ssDNA to make sufficient hybrid for cellular experiments would be very labour intensive and has a high risk of batch variability, in addition to the potential introduction of acrylamide contamination. Therefore, the use of other methods to purify larger amounts of ssDNA188 was investigated.

5.5.1. Fractionation of single- and double-stranded nucleic acids using silica spin columns

Commercially available silica spin-columns generally have a minimum size exclusion of approximately 50 bp, therefore theoretically could be appropriate for the binding of 188 nt ssDNA. The use of silica spin columns to fractionate large (> 1 kb) fragments of ssDNA from mixtures containing dsDNA has been reported previously (Beld *et al.*, 1996; Boom *et al.*, 1990; Borodina *et al.*, 2003). These methods exploit the differential binding of ss- and dsNA to silica when different binding buffers are used. Initially, the NA mixture is combined with a buffer containing a high concentration of the chaotrophic agent guanidium thiocyanate (GuSCN). When put over a silica column, only dsNA will bind and the ssNA is eluted. The ssNA-containing eluate is then mixed with a second binding buffer containing GuSCN and magnesium chloride ($MgCl_2$) and applied to a new silica column which it will then binds to. The two fractions are then eluted from the columns according to standard protocols. This method has also been applied to the fractionation of ss- and dsRNA (Beld *et al.*, 1996; Borodina *et al.*, 2003). Interestingly, it has been reported that RNA:DNA hybrids behave like dsNA (Beld *et al.*, 1996), although the size of the hybrid tested was not stated. Therefore, this method could be useful both for purification of the ssDNA and subsequently of the annealed RNA:DNA hybrid.

A mixture of undigested and λ exonuclease-digested dsDNA188 was processed according to the methods described by (Beld *et al.*, 1996; Borodina *et al.*, 2003). Despite extensive attempts at optimisation, binding of dsDNA to the first column was inefficient, but ssDNA completely failed to bind so the fraction was purely dsDNA (Figure 5.17). However, the $MgCl_2$-containing buffer did not seem to be specific for ssDNA. Consequently, the ssDNA fraction contained a mixture of ss- and dsDNA (Figure 5.17). In addition, the binding of the ssDNA to column was extremely inefficient and the majority of the ssDNA (estimated at least 90%) was lost. Therefore, this technique was not suitable for purifying ssDNA and the inefficient binding of the dsDNA meant it was not a practical option for recovering large amounts of dsNA.

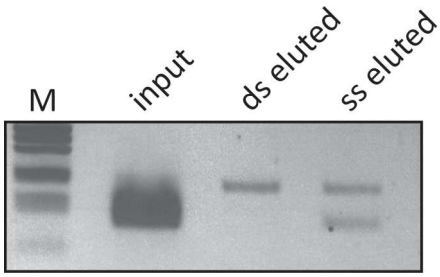


Figure 5.17. Separation of ds- and ssDNA using spin columns.
dsDNA188 was digested with λ exonuclease. An amount equivalent to 13.5 µg of undigested dsDNA was applied to a column as described in Chapter 2. The ds- and ssDNA fractions were eluted from the columns using TE. 10% of input and 15% of the fractions were electrophoresed on a 2% agarose gel containing ethidium bromide. Representative of three replicate experiments.

5.5.2. FPLC fractionation of single- and double-stranded DNA

Size exclusion chromatography using the Superdex 200 column in an FPLC system was successfully used to fractionate dsDNA from ssDNA of 40 nt in length (Section 5.2). The upper limit of this column, as determined by the manufacturers, is 200 bp DNA, however there is no defined limit for ssDNA.

To investigate the elution point for both ss- and dsDNA188, λ exonuclease-digested dsDNA188 was injected into the column. As shown in Figure 5.18 **a**, the main peak (peak 5) was at 8.7 - 9.0 ml, which is very close to the void volume of this column (approximately 7.5 ml, as shown in Figure 5.2). Native PAGE analysis of the fractions revealed that some dsDNA188 eluted in the void volume (peak 1, at 1.0 ml), whilst a mixture of ss- and dsDNA eluted just after the void volume (peak 4 at 8.4 ml and peak 5). Therefore, size exclusion chromatography using the Superdex 200 column did not successfully fractionate nucleic acids of this size. Ion exchange chromatography may be more successful; however a suitable column was not available to test this.

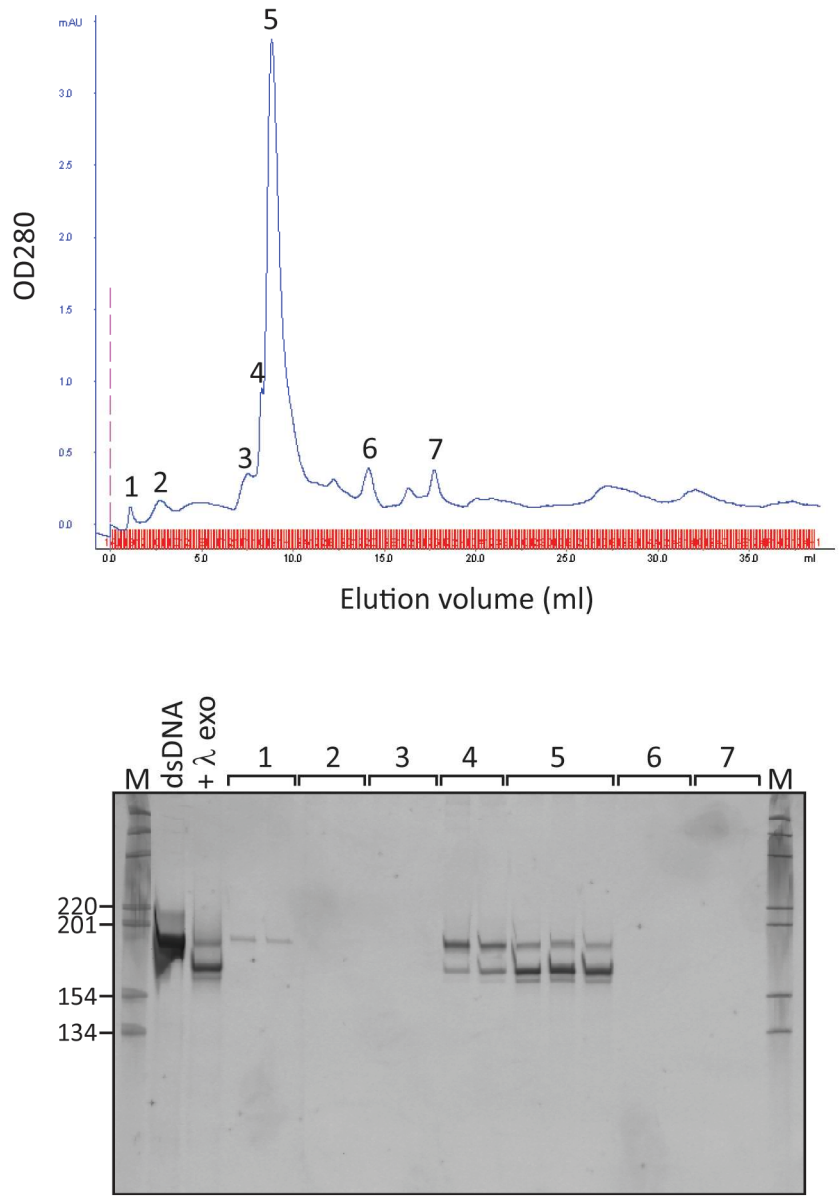


Figure 5.18. FPLC using a Superdex 200 column will not fractionate 188 nt nucleic acids. 3.5 µg of phosphorylated dsDNA188 was digested with λ exonuclease, ethanol precipitated and injected into a Superdex 200 column. Nucleic acid was eluted using TE. 20 µl of each fraction, together with 200 ng of undigested dsDNA188 and the equivalent volume of λ exonuclease digested dsDNA188 were electrophoresed on an 8% native polyacrylamide gel. 200 ng 1 kb ladder was included as a size marker (M). The gel was stained with Sybr Gold and imaged using a phosphorimager.

5.5.3. Separation of single- and double-stranded DNA using hydroxyapatite chromatography

The use of hydroxyapatite, a form of crystalline calcium phosphate, to fractionate nucleic acids is a well established technique (Bernardi, 1965). The positively charged calcium ions on the surface of hydroxyapatite can form non-specific interactions with the negatively charges phosphate groups in the backbones of nucleic acids. Both ss- and dsNA bind to hydroxyapatite when in a 50 mM sodium phosphate (pH 6.8) buffer (Sambrook and Russell, 2000). They are differentially eluted using an increasing gradient of phosphate buffers. The structure of ssNA results in fewer phosphate groups available to bind the hydroxyapatite and consequently they are eluted at lower buffer concentrations than dsNA.

In addition to the fractionation of ss- and dsDNA and RNA (Andrews-Pfannkoch *et al.*, 2010; Olgiati *et al.*, 1976), hydroxyapatite column chromatography has been successfully used to separate RNA:DNA hybrids from ssNA (Maxwell *et al.*, 1978; Siebke and Ekren, 1970), although short (36 bp) duplexes do not bind to hydroxyapatite very efficiently (Maxwell *et al.*, 1978). Therefore, this technique could be a useful method for the separation of both λ exonuclease digested ssDNA188 and a long RNA:DNA hybrid from their respective contaminating nucleic acids.

A 1.0 ml hydroxyapatite column was prepared as described in Chapter 2, and a mixture of digested and undigested dsDNA188 bound to the resin using 50 mM sodium phosphate at room temperature. A series of phosphate buffers of increasing concentrations from 100 mM to 400 mM were washed applied to the column. All of the input bound strongly to the hydroxyapatite, as nucleic acid was not present in either the non-bound fraction (NBF) or the 50 mM phosphate wash fraction (Figure 5.20). According to the manufacturer, ssDNA should elute at 140 – 160 mM and dsDNA at 340 – 360 mM from the batch of hydroxyapatite used for these experiments. ssDNA was eluted at 120 mM and 140 mM; however the fractions also contained some dsDNA (Figure 5.19). Between 200 – 250 mM a mixed population

of ss- and dsDNA is eluted, however after 300 mM, predominantly dsDNA is recovered. At concentrations below 120 mM and above 360 mM, no nucleic acid was recovered. Therefore, hydroxyapatite column chromatography failed to completely fractionate ss- and dsDNA. Hydroxyapatite chromatography is frequently performed at 60^{o}C, whereas these experiments were performed at room temperature (approximately 20^{o}C). Although absorption and elution rates are reported to be the same regardless of temperature (Sambrook and Russell, 2000), performing the chromatography at 60^{o}C may be a method of further optimising the fractionation ss- and dsDNA188.

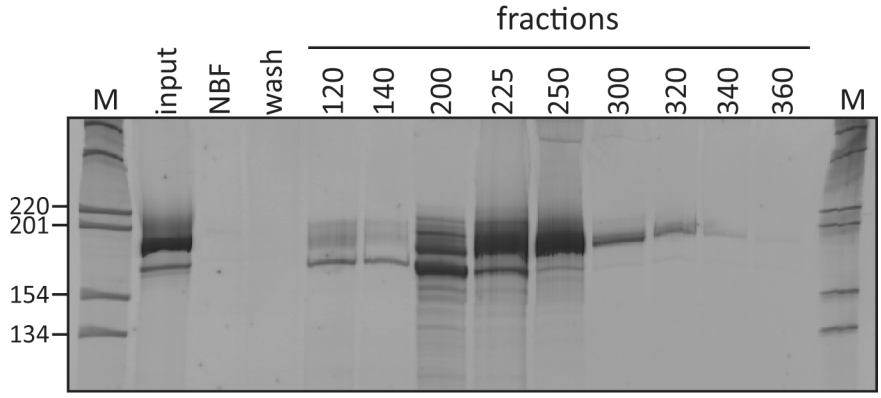


Figure 5.19. The fractionation of single- and double-stranded DNA using hydroxyapatite column chromatography.
Hydroxyapatite column chromatography was performed as described in Chapter 2. The ethanol precipitated fractions quantified using a NanoDrop. 300 ng of each fraction and 2% of the input mixture of ss- and dsDNA was electrophoresed on an 8% native polyacrylamide gel. 200 ng 1 kb ladder (M) was included as a size marker. The ssDNA band migrated faster than the dsDNA band. The gel was stained with Sybr Gold and imaged using a phosphorimager. Representative of two independent experiments.

5.5.4. Summary of alternative approaches to generating a long RNA:DNA hybrid

The production of a long (188 bp) RNA:DNA hybrid without utilising reverse transcription is complicated by the requirement to synthesise sufficient quantities of

pure ssDNA. The techniques discussed in this section revealed that λ exonuclease digestion of the forward strand of a PCR product was the most successful, and required the least manipulation of the nucleic acids. However, it was not possible to establish a method to fully purify the ssDNA from undigested dsDNA. Therefore the products of digestion would have to be excised from native gels to ensure only the correct strand is purified. Post-annealing, it is highly likely that the hybrid would need further purification, to remove any unannealed ssNA, as shown for the short RNA:DNA hybrids in Section 5.2. The techniques discussed in this section are also likely to be unsuitable for the separation of a 188 bp RNA:DNA hybrid from ssNA contaminants of the same length, therefore the hybrid would also require purification from native gels. This process would therefore require purification of each strand and the hybrid by gel extraction, and so a large quantity of starting material would be required to generate even a relatively small amount of product. RNA:DNA hybrids generated during this chapter were to be used to stimulate cells, therefore a large quantity of purified hybrid would be required to perform sufficient experiments. Consequently, R:D188 synthesised using this alternative method was not tested in cells during this thesis.

5.6. Additional RNA:DNA hybrids

RNA ligands for the cytosolic helicase RIG-I contain a 5′-ppp, which is recognised and strongly bound by RIG-I (Wang *et al.*, 2010a). RNA without a 5′-ppp does not stimulate RIG-I activity (Myong *et al.*, 2009). It is possible the innate immune signalling in response to RNA:DNA hybrids also requires the detection of a 5′-ppp on the RNA strand by the PRR.

An RNA strand produced by IVT will contain a 5′-ppp, therefore short RNA:DNA hybrids containing a 5′-ppp RNA were produced by combining the methods described for generating the other hybrids described in this chapter. A smaller region of the CMV promoter was amplified and used as a template for IVT. The resulting ssRNA was purified from a denaturing PAGE gel and annealed to a fully complementary 51 nt ODN. Initial PAGE analysis confirmed that a hybrid had been

formed by annealing and also revealed the presence of additional contaminating nucleic acids (Figure 5.20). Therefore, this hybrid would require further purification before used in cellular experiments. This hybrid is close to R:D60 in length, therefore the use of a Superdex 200 column to fractionate the different nucleic acid species would be expected to be successful. Although further purification and validation of this hybrid was not performed during this thesis, R:D51 illustrates that short hybrids containing 5′-ppp can be generated using this method, therefore providing an additional type of hybrid to be tested for immunostimulatory potential.

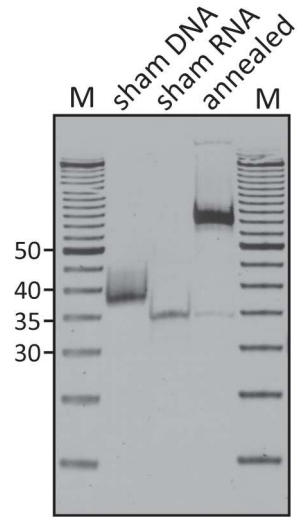


Figure 5.20. The annealing of a 51nt IVT RNA with a complementary ODN to generate a 51 bp RNA:DNA hybrid.
A region of the CMV promoter was amplified using a forward primer containing an 18 nt T7 polymerase promoter site and a reverse primer containing an *Eco*RI restriction site, digested with *Eco*RI, gel extracted and used as a template for IVT. RNA was extracted from a denaturing PAGE gel and then annealed to a complementary PAGE-purified DNA oligodeoxyribonucleotide as described in Section 5.2. Sham annealings omitting one of the strand were also performed: sham RNA (containing the DNA strand only) and sham DNA (containing the RNA strand only). 200 ng of annealed hybrid and the equivalent volume of each of the sham annealing were electrophoresed on a 15% native polyacrylamide gel, alongside 200 ng of 5 bp ladder as a size marker (M). The gel was stained with Sybr Gold and bands visualised using a UV transilluminator.

5.7. Discussion

In this chapter, the design, synthesis, purification and validation of synthetic RNA:DNA hybrids has been investigated. I have shown that whilst it is possible to generate large amounts of RNA:DNA hybrids, the removal of contaminating nucleic acid species is more challenging. However, removal of such contaminating species is essential before the hybrids can be tested for their ability to activate the innate immune response *in vitro*.

5.7.1. RNA:DNA hybrids generated by annealing complementary oligonucleotides

The annealing of complementary oligonucleotides is the most frequently used method to generate RNA:DNA hybrids. However, purification of hybrids post-annealing is rarely performed. This is most probably because the annealed product is considered sufficiently pure for downstream applications, including RNase H activity assays (Ohtani *et al.*, 2008) and co-crystallisation experiments (Nowotny *et al.*, 2008). Whilst annealed RNA:DNA hybrids have not previously been used to investigate immune responses *in vitro*, the use of short immunostimulatory dsDNA and dsRNA generated by the annealing of complementary ODN or ORNs has been reported in a number of studies, for example (Karayel *et al.*, 2009; Schlee *et al.*, 2009b; Stetson and Medzhitov, 2006). In all cases, purification of the annealed duplexes prior to downstream experiments was not reported.

Native PAGE analysis of RNA:DNA hybrids generated by annealing complementary ORN and ODNs revealed the presence of unannealed ssNA in the preparations, when stained with the highly sensitive nucleic acid stain Sybr Gold (Figure 5.1 **b**). Furthermore, the annealing of certain complementary oligonucleotides resulted in the formation of higher molecular weight RNA:DNA hybrid-containing species (Figure 5.1 **b** and **c**; Chapter 4, Figure 4.6 **b**), most likely as a consequence of the formation of higher ordered structures upon annealing, or the use of insufficiently purified oligonucleotides. It is therefore conceivable that the use of unpurified hybrids generated using this method to stimulate an immune response may well cause

spurious results. This could arise as a direct consequence of the presence of contaminating nucleic acid species, in addition to the problems associated with accurately quantifying a hybrid contaminated with ssNA (as discussed in Chapter 4). However, the generation of short RNA:DNA hybrids using this method has many advantages: it is a simple and rapid method of producing large quantities of hybrids, which can be generated using ODNs and ORNs with a number of chemical modifications of any sequence. Therefore, purification of the correct hybrids from other nucleic acids using FPLC size exclusion chromatography was established.

5.7.2. FPLC purification of R:D60

The use of FPLC provided a rapid method of fractionating large quantities (up to 150 μg input per run was used during this thesis) of RNA:DNA hybrids from contaminants. The use of nuclease-free reagents and thorough washing of the column ensured that samples were protected from degradation and external contaminants. As shown in Section 5.2, excellent separation of 40 bp and 60 bp duplexes from ssNA was achieved using this method. The process was found to be more efficient than PAGE purification and had the additional advantage of being able to fractionate large quantities of hybrid at the same time, therefore reducing the possibility batch variability, which is preferable for screening the immune response to the same hybrid over a series of experiments.

R:D60 fractionated using FPLC was generally recovered completely free from contaminants (Figure 5.4). Quality controls experiments were performed to monitor the maximum amount of contamination (Figure 5.7). Unannealed ssDNA was present in very few purified preparations, and could be attributed to technical problems (collection of the incorrect fractions and failure to sufficiently clean the Superdex column). These issues were resolved by re-running the samples over the column. Nevertheless, the quality control gel shown in Figure 5.7 allows a maximal amount of ssDNA contamination to be estimated, which is approximately 6% for this hybrid batch. Using this method, it was possible to estimate that the R:D60 used in the downstream experiments investigating stimulation of innate immunity described

in this thesis (Chapter 6) contained less than 1.6% ssNA contamination (as 1.6 ng is the limit of Sybr Gold detection, Figure 5.7).

A limitation of this technique was revealed when used to fractionate nucleic acids of other sizes. Although complete separation of 18 nt ssDNA and the 18 bp hybrid was achieved, the purity of the recovered hybrid-containing fractions of R:D12 and R:D20 was lower (Figure 5.5). Furthermore, 188 bp dsDNA and 188 nt ssDNA could not be resolved (Figure 5.18). Therefore, it can be concluded that successful fractionation of ssNA and dsNA using size exclusion chromatography with a Superdex 200 column is dependent upon the both length of the nucleic acid and the base sequence, which influences the migration and secondary structure of ssNA. It may be possible to achieve better fractionation of these other samples using different columns which may have more appropriate size selection ranges.

5.7.3. Synthesis of RNA:DNA hybrids containing 5′-ppp RNA

The overall aim of this thesis was to determine if RNA:DNA hybrids can stimulate innate immunity *in vitro* as a model of what may occur *in vivo* in AGS patient cells. Therefore, the use of physiologically relevant hybrids would be preferential, although as discussed previously during this thesis, many studies performing *in vitro* characterisation of innate immune signalling pathways have used synthetic nucleic acid analogues and ODNs.

The production of a much longer hybrid containing virally-derived sequence was established, using reverse transcription of a 188 nt ssRNA generated by IVT (Figure 5.8). Although this method produced a hybrid-containing species of the correct size which could be hydrolysed by RNase H2, it was not possible to exclude ssNA contamination from this hybrid (Figure 5.10). A number of methods to synthesise sufficient quantities of pure 188 nt ssDNA were investigated, with the aim of generating a 188 bp RNA:DNA hybrid by annealing the two separate strands. It was not possible to obtain sufficiently purified ssDNA during this thesis, however the use of an alternative method of chromatography, for example, ion exchange

chromatography, may be more successful in fractionating λ exonuclease digested DNA.

Finally, a 51 bp RNA:DNA hybrid with a 5′-ppp ssRNA strand was generated (Figure 5.20). Assuming the hybrid can be successfully purified away from contaminating nucleic acid species using FPLC, this provides an additional type of hybrid to be tested for the ability to activate the PRRs of the innate immune system. Transcription of certain endogenous retroelements by RNA pol III generates ssRNA with a 5′-ppp (Boeke, 2003). Additionally, the short RNA primers synthesised by the primase complex during DNA replication also have a 5′-ppp (Wilson and Sugino, 1985). Therefore, it possible that endogenous immunostimulatory RNA:DNA hybrids will contain a 5′-ppp, which may be vital to the detection of hybrids by PRRs. The production of hybrids similar to R:D51 would allow the generation of hybrids of various sizes (< 100 bp) to be produced from any sequence that can be used as a template for IVT. The inclusion of a cap analogue during transcription, or the treatment of the RNA with a phosphatase would allow for the comparison of the immunostimulatory potential of the same hybrid with and without a 5′-ppp, similar to studies investigating the ligand for RIG-I (Hornung *et al.*, 2006; Pichlmair *et al.*, 2006; Rehwinkel *et al.*, 2010).

In summary, I have described the design and synthesis of seven RNA:DNA hybrids with different physiochemical properties. All hybrids required purification to remove any contaminating nucleic acid species, most notably ssNA which is likely to be immunogenic. A 60 bp hybrid consisting of a repetitive GU-repeating RNA sequence generated from unmodified PD-containing oligonucleotides was successfully purified. Further characterisation of the hybrid using detection by the S9.6 antibody combined with RNase H digestion confirmed that FPLC-fractionated R:D60 is a homogeneous RNA:DNA hybrid. Chapter 6 describes details the testing of several RNA:DNA hybrids generated during this chapter for ability to stimulate innate immune signalling *in vitro*.

Chapter 6. The stimulation of innate immunity *in vitro* by RNA:DNA hybrids

Loss-of-function mutations in the *RNASEH2* genes are hypothesised to lead to an accumulation of RNA:DNA hybrid substrates in cells. The primary aim of this thesis was to determine if these substrates are able to stimulate an innate immune response. Synthetic RNA:DNA hybrids generated in Chapter 5 were used to investigate this *in vitro* via the stimulation of different cell types. Initially, the 188 bp 5′-ppp-containing hybrid, R:D188, was to be the test hybrid as a long hybrid formed by reverse transcription was considered to be a more physiologically relevant hybrid. However, failure to validate the precise nature of this hybrid, in addition to the problems with purification of the hybrid discussed in Chapter 5, led to the examination of the immunostimulatory potential of shorter hybrids.

6.1. Characterisation of reagents for the transfection of RNA:DNA hybrids into cells

Studies using synthetic nucleic acids to stimulate activation of an innate immune response *in vitro* frequently utilise cationic lipid-mediated transfection to deliver the nucleic acid into a wide variety of cell types (Ablasser *et al.*, 2009; Boczkowski *et al.*, 1996; Burckstummer *et al.*, 2009; Chiu *et al.*, 2009; Diebold *et al.*, 2003; Fernandes-Alnemri *et al.*, 2009; Heil *et al.*, 2004; Hornung *et al.*, 2009; Hornung *et al.*, 2006; Ishii *et al.*, 2006; Roberts *et al.*, 2009; Stetson and Medzhitov, 2006). In addition to facilitating cellular uptake, the complexing of nucleic acids to cationic lipids is thought to protect from degradation by extracellular nucleases (Heil *et al.*, 2004). In the absence of transfection, the immunostimulatory potential of most synthetic oligonucleotides is negated (see references listed above).

To investigate if RNA:DNA hybrids required transfection for entry into cells, an 18 bp RNA:DNA labelled with fluorescein on the RNA strand hybrid (*R:D18, where * denotes a 5′-fluorescein) was added to the culture medium of transformed cell lines

(human epithelial HeLa cells and murine L929 fibroblasts) and primary MEFs. After a series of time points (10 min – 24 h), the cells were analysed by microscopy and flow cytometry. Fluorescent signal was not detected in any of the cell lines over any time point, indicating that the hybrids had failed to enter the cell and transfection would be required (Figure 6.1 **b**).

The ability of various cationic lipid transfection reagents to facilitate entry of 18 bp RNA:DNA hybrids into cells was assessed using *R:D18. Each reagent successfully delivered *R:D18 into the cytosol (Figure 6.1 **a**). However, the efficiency of transfection and distribution of signal throughout the cytosol varied between reagents, with the diffuse signal seen in cells transfected with using Lipofectamine 2000 considered to be the most physiologically relevant. Therefore, Lipofectamine 2000 was chosen as the transfection reagent for subsequent experiments. After 4 h, *R:D18 complexed with Lipofectamine 2000 in 99% of cells (Figure 6.1. **b** and **c**). Although this was only shown for using 18 bp RNA:DNA hybrid, the published requirement for transfection of other nucleic acids makes it likely that longer RNA:DNA hybrids will also require transfection to enter cells.

Lipofectamine 2000 is frequently used in studies to stimulate an immune response to nucleic acid ligands in a variety of cells types including transformed cell lines, and primary MEFs, macrophages, monocytes and DCs (see references listed above). Importantly, non-specific cytokine production by immune cells is not induced by Lipofectamine 2000 (Diebold *et al.*, 2003).

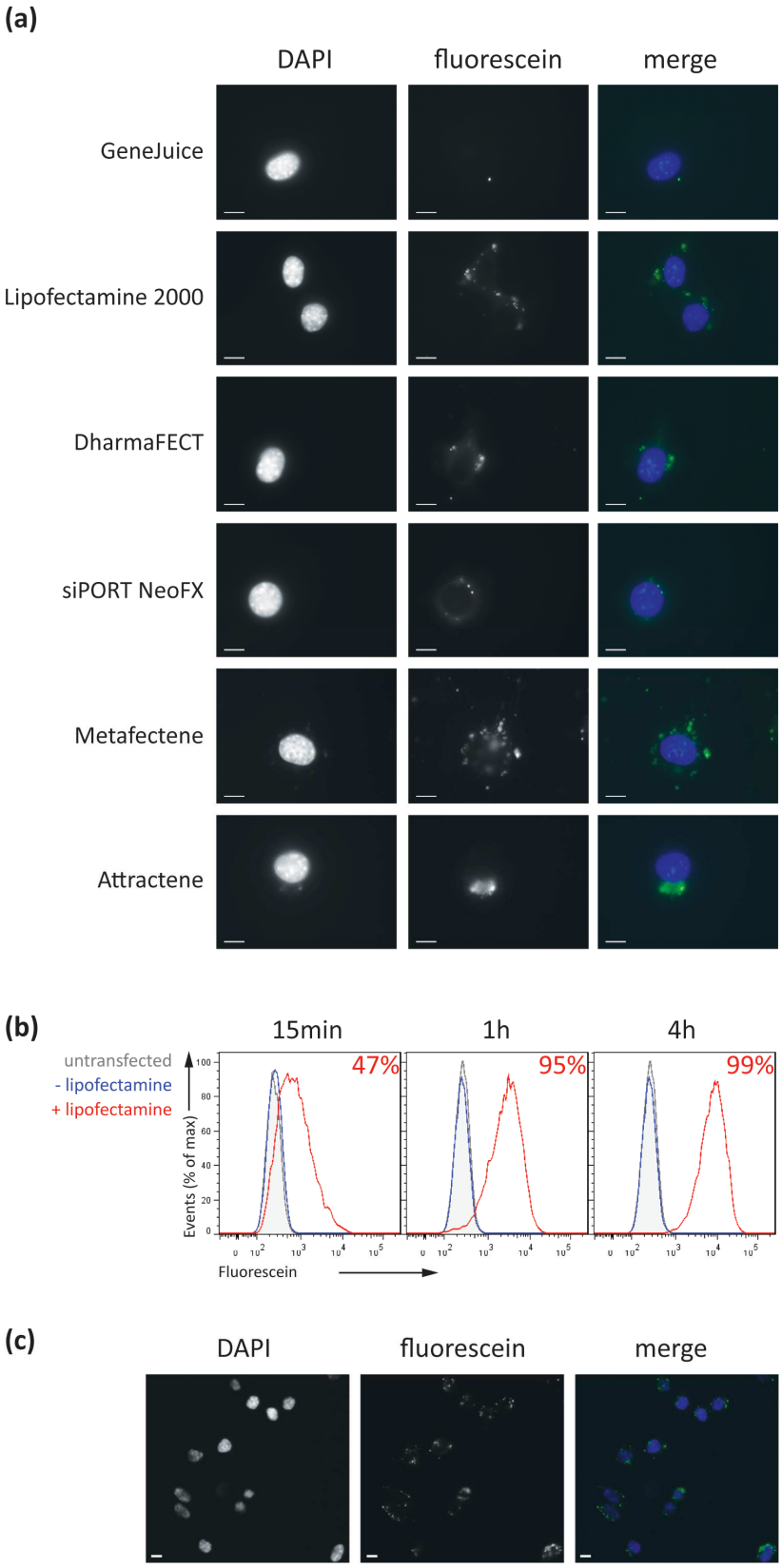


Figure 6.1. Characterisation of reagents for the transfection of R:D18 into L929 cells. Semi-confluent L929 cells were transfected with 4.4 µg/ml *R:D18 according to the manufacturer's instructions for each transfection reagent. After 4 h at 37°C, transfection

medium was removed, and the cells fixed and stained with DAPI. **(a)** All transfection agents delivered *R:D18 to the cytosol. Cells shown are representative of duplicate slides from two independent experiments using L929 cells. Scale bars are 10 μm, images taken at a magnification of 100 x. Similar results were also seen in HeLa cells and primary MEFs and when different concentrations of hybrid were transfected (1 – 4.4 μg/ml). **(b)** Efficiency of *R:D18 transfection was analysed using flow cytometry. Cells were transfected with 1 μg/ml *R:D18 for the indicated time. After 4 h, 99% of cells contained *R:D18, as shown in the field of view in **(c)** (scale bar = 10 μm, taken at a magnification of 20 x).

6.2. R:D188 stimulates IFN production in transformed cell lines

In initial experiments, two transformed cell lines were stimulated with both short and long RNA:DNA hybrids. Both HeLa and L929 cells are known to produce type I IFN in response to transfection with poly (I:C) (Milhaud *et al.*, 1992) and infection with RNA viruses (van Pesch *et al.*, 2004; Wathelet *et al.*, 1992). Unstimulated HeLa cells express mRNA for most TLRs, although expression of TLR8 is weak (Kurt-Jones *et al.*, 2004; Nishimura and Naito, 2005). In addition, they can utilise the RNA helicases RIG-I/MDA5 (Choi *et al.*, 2009), as can L929 cells (Yoneyama *et al.*, 2004).

Addition of poly (I:C) to culture medium stimulates type I IFN expression via TLR3 (Oshiumi *et al.*, 2003), whereas transfection of poly (I:C) stimulates IFN induction independent of TLR3, via MDA5 (Andrejeva *et al.*, 2004; Gitlin *et al.*, 2006; Yoneyama *et al.*, 2004). In the following experiments, poly (I:C) used as a positive control for IFN induction was transfected into the cells, since RNA:DNA hybrids also required transfection for successful entry into the cell.

Transfection of short hybrids into HeLa cells did not stimulate an IFN response (as shown for R:D18 in Figure 6.2 **a**). In contrast, transfection of R:D188 produced by reverse transcription (Chapter 5, Section 5.3) induced a type I IFN response, as shown by an significant increase in the mRNA levels of the genes encoding IFN-β (Student's *t*-test, $p = 0.0056$) and IRF7 ($p = 0.0031$) compared to Lipofectamine

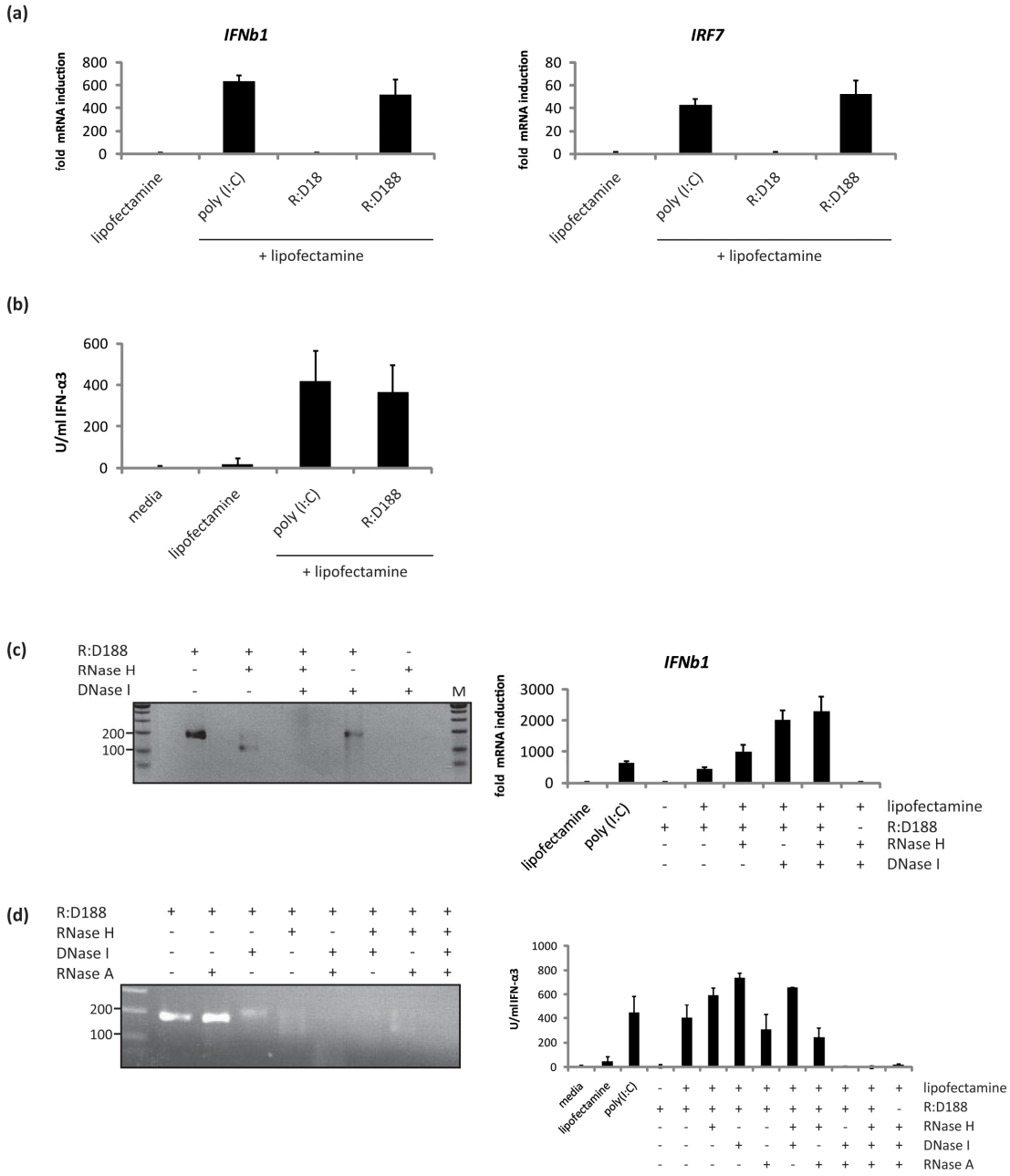


Figure 6.2. R:D188 stimulates type I IFN production in transformed cell lines.
Cells were transfected with 25 µg/ml each of poly (I:C), R:D18 and R:D188 using Lipofectamine 2000 for 18 h. **(a)** HeLa cells upregulated expression of the genes encoding IFN-β and IRF7 in response to stimulation with R:D188. Data are shown as fold increase in expression from medium only controls, from three independent experiments ± s.d. **(b)** L929 cells secreted IFN-α in response to R:D188. Data are from three independent experiments ± s.e.m. **(c)** Nuclease digestion of R:D188 did not abolish the IFN-β response in HeLa cells. Data shown are from one experiment ± s.d. of triplicate qRT-PCR wells. **(d)** Nuclease digestion also failed to abolish IFN-α secretion by L929s in response to R:D188. Data shown are from three independent experiments ± s.e.m.

only controls (Figure 6.2 **a**). Similarly, transfected R:D188 robustly induced IFN-α secretion by L929 cells (Figure 6.2 **b**). However, in the absence of Lipofectamine 2000, R:D188 failed to induce a type I IFN response (Figure 6.2 **c** and **d**).

To establish if nuclease digestion would abolish the response of the cell lines to the long hybrid, R:D188 was digested with RNase H and DNase I. Agarose gel electrophoresis of the digested products showed that RNase H treatment hydrolysed the RNA strand of the hybrid, resulting in a band at approximately 150 bp (Figure 6.2 **c** and **d**). The migration of this strand and its digestion by DNase I suggested that this was the ssDNA strand of the hybrid, as confirmed by Figure 5.10. DNase I treatment of R:D188 also resulted in some degradation of the hybrid, which suggested non-specific activity of the enzyme (as discussed in Chapter 4). Alternatively, there could have been ssDNA contamination of R:D188; however this wasn't evident from agarose gel analysis of the hybrid.

The nuclease treated hybrids were transfected into cells. *IFNb1* expression was increased in the nuclease digested samples, and was induced most strongly when R:D188 was digested with both RNase H and DNase I (Figure 6.2 **c**). This suggested that the short polynucleotide products of nuclease digestion, which are not evident on an agarose gel, were highly immunostimulatory in HeLa cells. Similarly, nuclease digestion did not abolish IFN-α production by L929s in response to R:D188 (Figure 6.2 **d**). RNase A treatment in a high salt buffer, which did not degrade the hybrid as seen on the gel, did slightly reduce IFN-α production ($p = 0.572$), which suggested there may have been some contaminating ssRNA in R:D188. Alternatively, there may have been regions of ssRNA in the hybrid, for example, if looping of the RNA strand occurred during reverse transcription.

Together, these results demonstrated that in the HeLa and L929 cell lines, a long RNA:DNA hybrid with a 5′-ppp RNA strand can stimulate a type I IFN response when introduced into cells using a transfection reagent. However, nuclease digestion data, combined with the analysis of the hybrid described in Chapter 5, meant it was

not possible to verify that this response is due to an RNA:DNA hybrid and not contaminating nucleic acid species.

Short hybrids did not stimulate type I IFN production in HeLa and L929 cells. The methods used to produce these hybrids, including their successful purification from unannealed ssNA by FPLC, as detailed in Chapter 5, suggests they are more likely to be free from contaminating nucleic acids than R:D188. It is known that transformed cell lines lack certain PRR signalling pathways, such as the ISD response pathway (the response to cytosolic phosphodiester DNA) (Stetson *et al.*, 2008). Therefore the immunostimulatory capacity of short RNA:DNA hybrids was analysed using primary cells, including MEFs, and more appropriately, cells of the immune system.

6.3. Short RNA:DNA hybrids do not induce expression of type I IFNs in primary MEFs

The immunostimulatory potential of short RNA:DNA hybrids was initially tested in non-immune cells. Primary MEFs express all TLRs (Kurt-Jones *et al.*, 2004) but can also sense nucleic acids via cytosolic receptors. After infection with ssRNA viruses, MEFs produce type I IFN via a RIG-I-dependent mechanism (Kato *et al.*, 2005), whilst the B-form dsDNA poly (dA:dT) induces expression of the gene encoding the cytosolic DNA sensor DAI (Takaoka *et al.*, 2007).

Primary MEFs derived from wild type mouse embryos were established and their ability to upregulate type I IFN expression in response to stimulation with TLR ligands was tested. The control TLR ligands poly (I:C) and CpG ODN (Type A) potently stimulated expression of *Ifna1* and *Ifnb1* (Figure 6.3 **a**).

To investigate if type I IFN expression was upregulated in response to RNA:DNA hybrids, a range of hybrids were transfected into MEF cultures. As in the transformed cell lines, the long hybrid R:D188 strongly induced type I IFN expression (Figure 6.3 **b**). However, the short hybrids R:D12, R:D18, R:D20 and R:D60 and the commercially available hybrid poly (rA:dT) did not stimulate

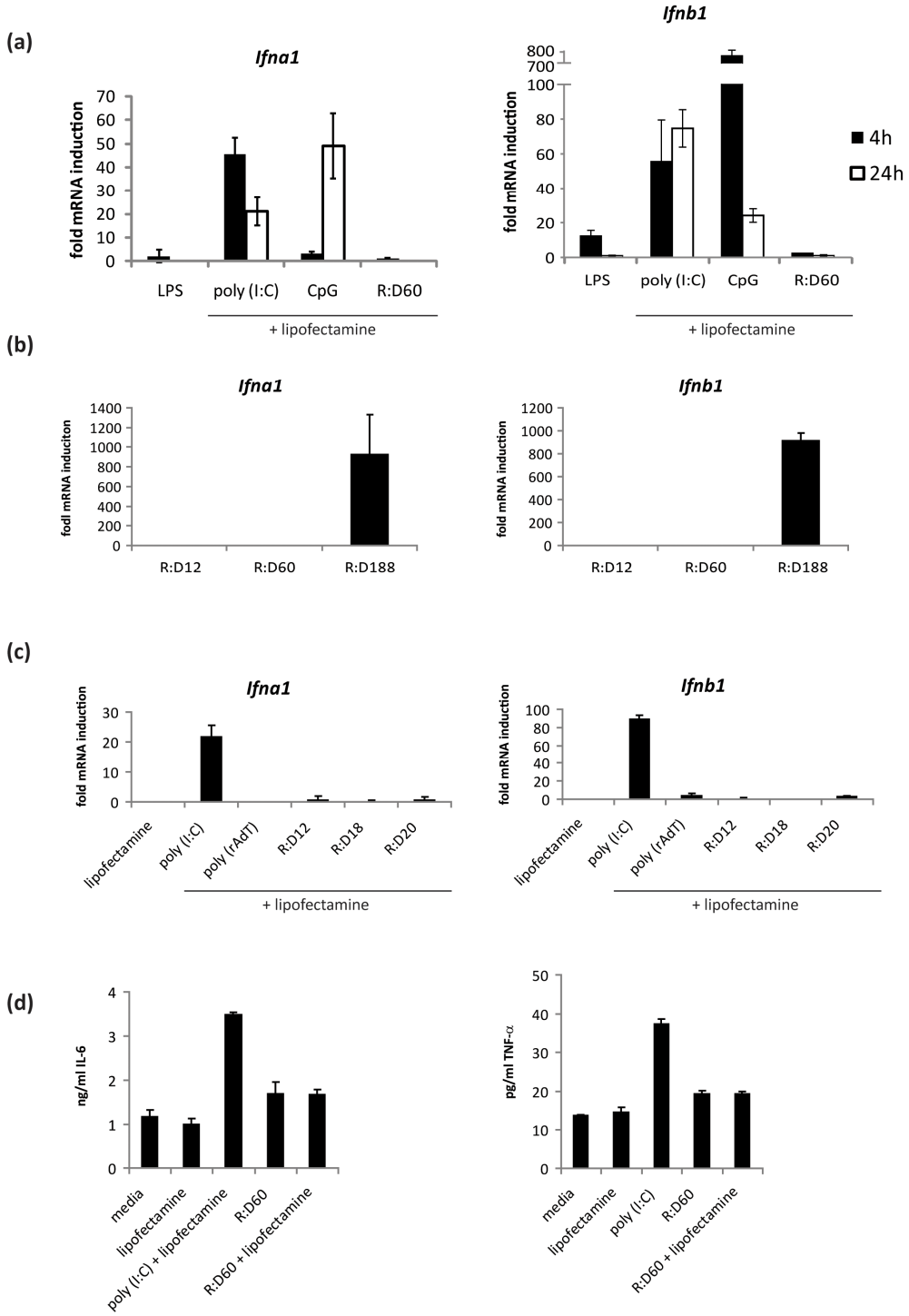


Figure 6.3. Expression of type I IFNs is not induced in primary MEFs following transfection with short RNA:DNA hybrids.

Semi-confluent early passage (< 3 passages) primary MEFs were transfected with TLR ligands and hybrids. None of the hybrids used in these experiments were FPLC-purified. **(a)** Type I IFN gene expression in MEFs stimulated with 200 ng/ml LPS, 5 µg/ml CpG ODN, 25 µg/ml poly (I:C) and 25 µg/ml R:D60 for the indicated timepoints. Data are represented as normalised fold change relative to Lipofectamine alone, representative of 3 independent experiments ± s.d. of triplicate qRT-PCR wells. **(b)** MEFs were transfected with 25 µg/ml

each of R:D12, R:D60 and R:188 for 18 h before IFN gene expression was analysed by qRT-PCR. Data are represented as normalised fold mRNA induction from Lipofectamine only controls ± s.d. of triplicate qRT-PCR wells. **(c)** *Ifna1* and *Ifnb1* expression in MEFs transfected with 25 µg/ml of poly (I:C) and each RNA:DNA hybrid for 18 h. Data are represented as normalised fold mRNA induction from a medium only control, representative of duplicate experiments ± s.d. of triplicate qRT-PCR wells. **(d)** IL-6 and TNF-α production by MEFs transfected with 25 µg/ml of poly (I:C) and R:D60 in the presence and absence of Lipofectamine for 18 h. Data are from two independent experiments ± s.d.

expression of *Ifna1* and *Ifnb1* (Figure 6.3 **a**, **b** and **c**) or production of the pro-inflammatory cytokines IL-6 and TNF-α. Furthermore, pretreating MEFs with IFN-α before transfection of the hybrids also failed to induce type I IFN expression. Therefore, it can be concluded that these short RNA:DNA hybrids could not trigger an IFN response in primary MEFs.

6.4. Short RNA:DNA hybrids do not stimulate an immune response in primary macrophages

To investigate whether short RNA:DNA hybrids can stimulate cytokine production in antigen-presenting cells of the immune system, primary bone-marrow derived macrophages were generated. In response to stimulation with DNA, including CpG ODN and poly (dA:dT), macrophages produce pro-inflammatory cytokines and type I IFNs (reviewed by Vilaysane and Muruve, 2009). Similarly to MEFs, macrophages respond to RNA viral infection via cytosolic RLR pathways, leading to the production of type I IFN and pro-inflammatory cytokines (Kumar *et al.*, 2006).

Bone-marrow macrophage cultures produced IFN-α in response to transfection with poly (I:C) (Figure 6.4 **a**). R:D188 also induced the production of small amounts of IFN-α, however R:D60 did not (Figure 6.4 **a**). The expression of *Ifnb1* was upregulated in response to transfection with nucleic acid PRR ligands (Figure 6.4 **b**). However, no increase in expression was detected following transfection with either R:D12, R:D18, R:D20 or R:D60 (Figure 6.4 **b**). The single-stranded oligonucleotides used to produce R:D60 also failed to induce *Ifnb1* expression (Figure 6.4 **b**). This

suggested that short RNA:DNA hybrids and ssNA cannot trigger a type I IFN response in macrophages.

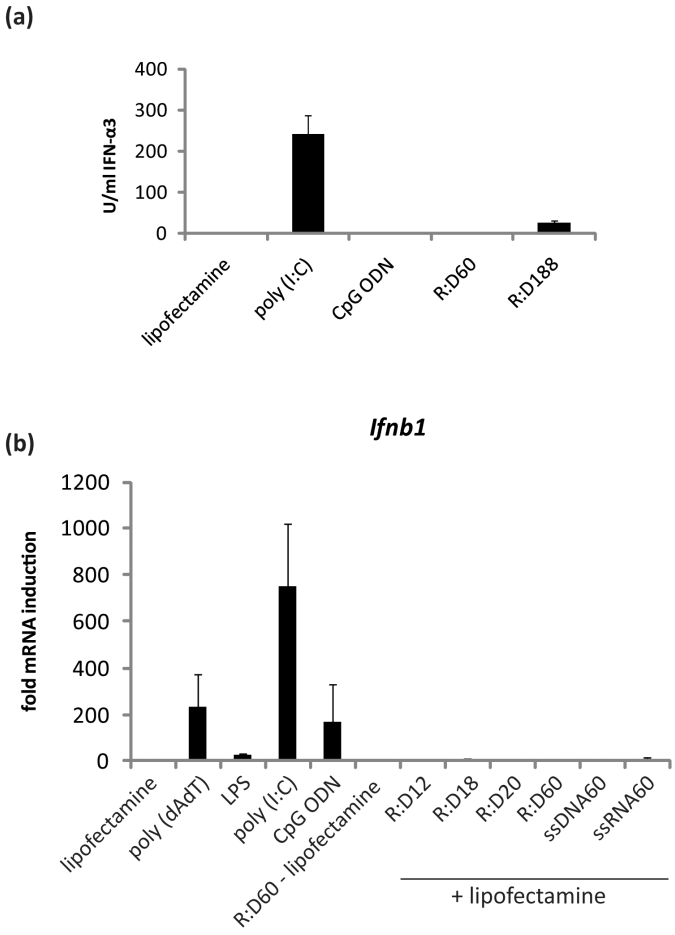


Figure 6.4. Short RNA:DNA hybrids do not induce a type I IFN response in macrophages. **(a)** Macrophages were transfected with 25 µg/ml poly (I:C), 10 µg/ml CpG ODN (Type A) and 25 µg/ml of each hybrid for 18 h before IFN-α secretion was determined by ELISA. Data is from one experiment performed in triplicate ± s.d. **(b)** Macrophages were stimulated with 100 ng/ml LPS or transfected with 1 µg/ml each of poly (dA:dT), poly (I:C) or CpG ODN. R:D60 was added to to the medium or transfected at 1 µg/ml. The hybrids R:D12, R:D18 and R:D20 were transfected at 1 µg/ml and the ssNA controls for R:D60 were transfected at 500 ng/ml (the equivalent amount of each that would be present in 1 µg of R:D60). After 18 h, cytokine expression was determined by qRT-PCR. No *Ifna1* mRNA was detected in any of the samples. Data shown are from two independent experiments ± s.d.

Macrophages also produce pro-inflammatory cytokines in response to nucleic acid ligands. Levels of IL-6 production following treatment with different concentrations of R:D60 and its single-stranded components ssDNA60 and ssRNA60 was measured by ELISA. Transfected R:D60 failed to stimulate IL-6 production as did ssRNA60, however a high concentration of ssDNA60 induced a large amount of IL-6 production (25 μg/ml ssDNA60: Student's *t*-test, $p = 0.0095$) (Figure 6.5).

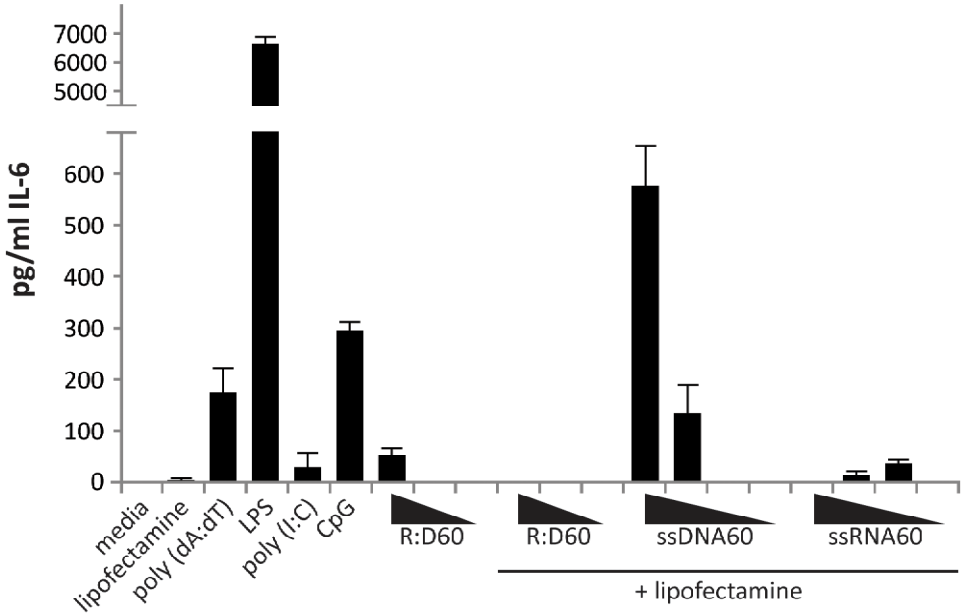


Figure 6.5. Transfected R:D60 does not induce pro-inflammatory cytokine production by macrophages.
Bone-marrow macrophage cultures were stimulated with 100 ng/ml LPS, or transfected with 1 μg/ml each of poly (dA:dT), poly (I:C) or CpG ODN. R:D60 was added to the medium or transfected at decreasing concentration (25, 10, 1 μg/ml). ssDNA60 and ssRNA60 were transfected at decreasing concentrations (25, 10, 1, 0.5 μg/ml). After 18 h, supernantants were harvested and IL-6 production determined by ELISA. Data shown are from two independent experiments ± s.d.

The expression of activation markers by macrophages transfected with R:D60 was analysed using flow cytometry. There was no significant upregulation of CD86 (Figure 6.6) or MHC Class II in cells treated with R:D60, either in the presence or

absence of Lipofectamine. Together, these data show that bone marrow-derived macrophages do not produce type I IFNs in response to short RNA:DNA hybrids, although the addition of high concentrations of R:D60 to culture medium can stimulate the production of low levels of pro-inflammatory cytokines in the absence of upregulation of activation markers ($p = 0.054$).

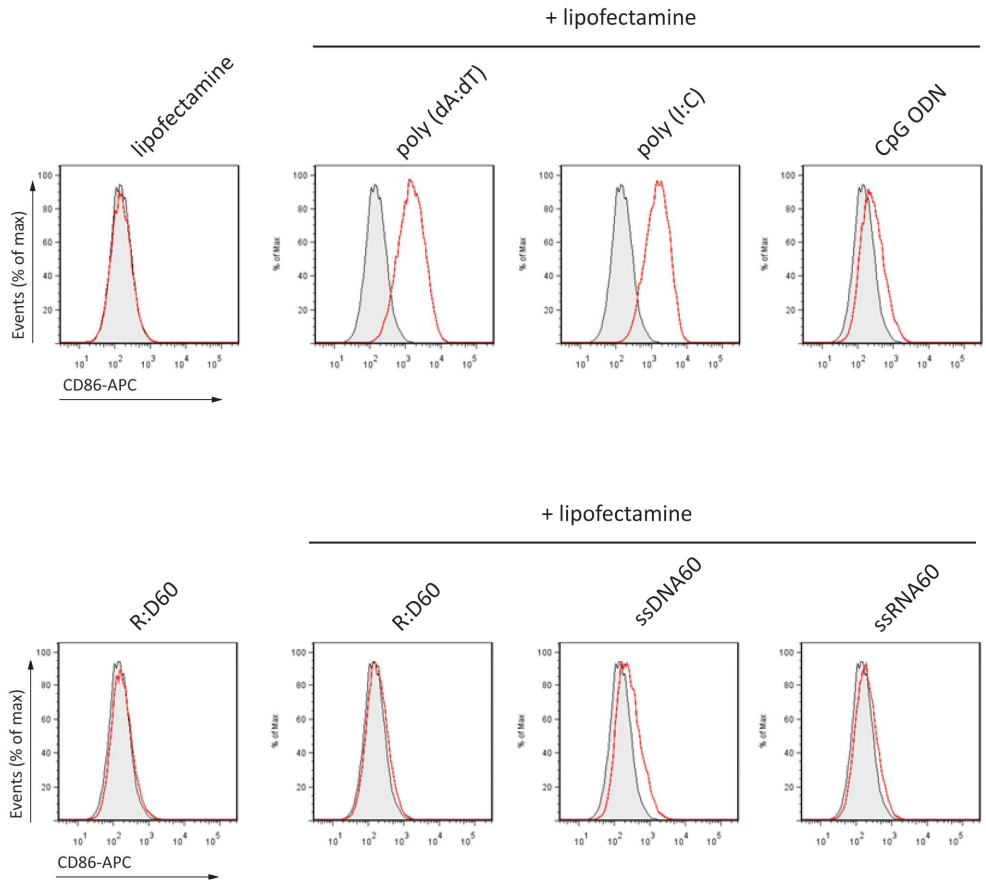


Figure 6.6. CD86 expression is not upregulated by macrophages transfected with R:D60. Bone-marrow macrophages were transfected with 1 µg/ml each of poly (dA:dT), poly (I:C) or CpG ODN. 25 µg/ml R:D60 was added to the medium or transfected. 25 µg/ml ssDNA60 and ssRNA60 were also transfected into cells. After 18 h, cell surface expression of CD86 was analysed by flow cytometry. Medium only controls are shown in grey, sample is in red. Data shown are representative of two independent experiments. These cells also failed to upregulate MHC Class II expression in response to R:D60 and ssNA.

6.5. Short RNA:DNA hybrids do not stimulate an immune response in GM-CSF DCs

DCs can be generated from bone marrow precursors *in vitro* using granulocyte macrophage colony-stimulating factor (GM-CSF) (Lutz *et al.*, 1999). GM-CSF DCs contain both mature and immature cells, as determined by surface marker expression (Lutz *et al.*, 1999) and are classed as conventional DCs (cDCs, also known as myeloid DCs (mDCs)). GM-CSF cultures do not contain plasmacytoid DCs (pDCs) (Brawand *et al.*, 2002).

GM-CSF DCs have been shown to produce type I IFNs and pro-inflammatory cytokines in response to stimulation with DNA, particularly the TLR9 ligand CpG ODN (Gray *et al.*, 2007; Yoshinaga *et al.*, 2007). They produce IFN-α in response to transfection with poly (I:C) and viral dsRNA via activation of cytosolic RLR-mediated pathways (Diebold *et al.*, 2003; Kato *et al.*, 2005; Kato *et al.*, 2006), but can also produce type I IFN in response to poly (I:C) when it is added to the culture medium, via TLR3.

GM-CSF DCs were stimulated with the TLR ligands CpG ODN, R848 and poly (I:C). R:D60 was included in the presence and absence of Lipofectamine. As expected, GM-DCs produced IFN-α in response to poly (I:C) and IL-6/IL-12p40 in response to R848 (Figure 6.7). R:D60 failed to stimulate a cytokine response in these cells, in the presence or absence of Lipofectamine. There was no phenotypic activation of GM-CSF DCs in response to the hybrid, evident by the absence of an increase of CD40, CD80, CD86 and MHC Class II expression on the cell surface (Figure 6.8). Therefore, R:D60 was unable to stimulate an innate immune response in GM-CSF DCs.

A second DC subset, plasmacytoid DCs (pDCs), are known to produce vast amounts of type I IFN in response to nucleic acids, particularly during viral infection (Asselin-Paturel *et al.*, 2001; Cella *et al.*, 1999). Therefore, the ability of R:D60 to trigger an IFN response in DC cultures containing pDCs was investigated.

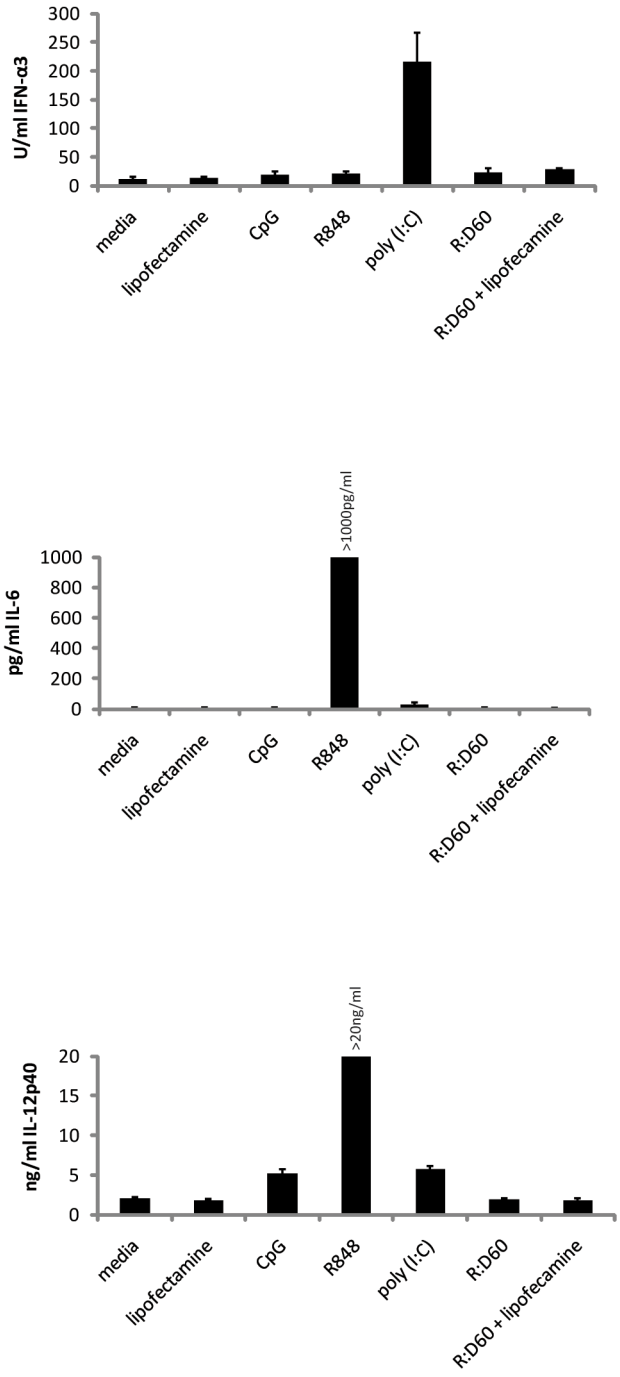


Figure 6.7. R:D60 does not stimulate cytokine production in GM-CSF DCs.
GM-CSF DCs were stimulated with 500 ng/ml CpG ODN, 1 µg/ml R848 and 1 µg/ml poly (I:C). R:D60 was added to the medium or transfected at a concentration of 1 µg/ml. After 18 h, supernatants were harvested and the levels of IFN-α, IL-6 and IL-12p40 determined by ELISA. Data shown are the mean from triplicate samples $\pm$ s.d.

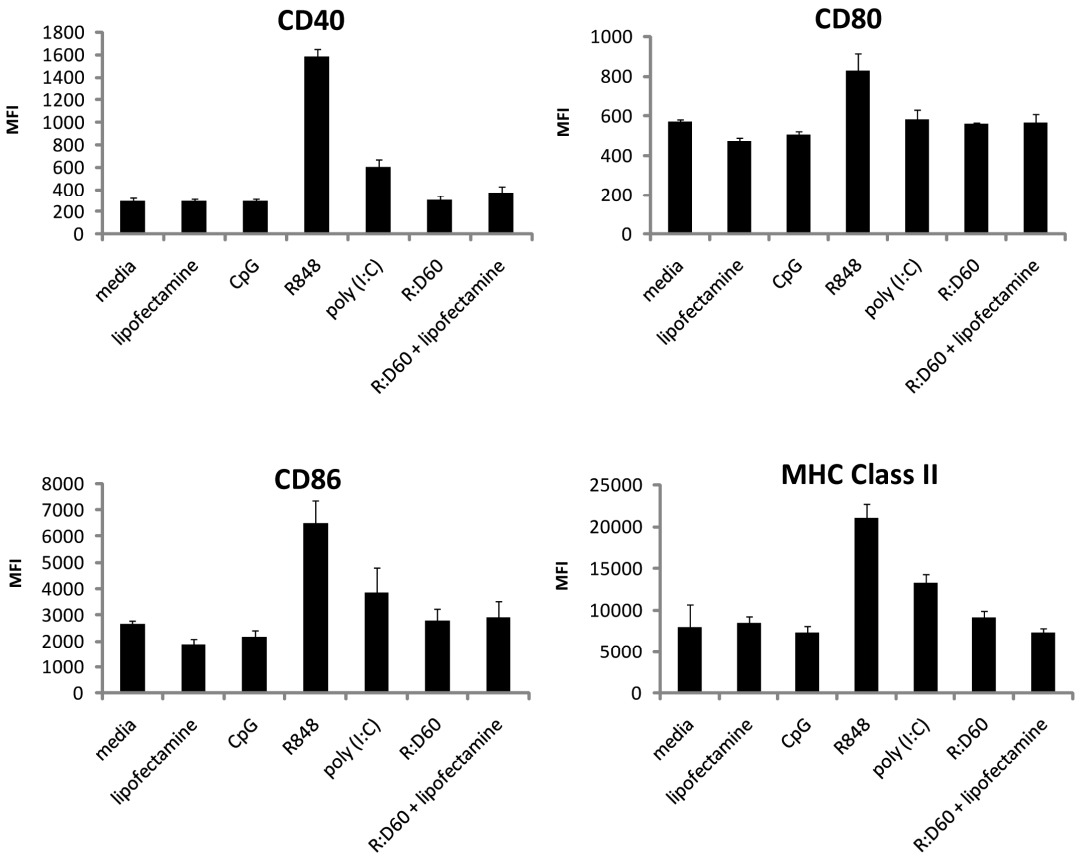


Figure 6.8. R:D60 does not stimulate the phenotypic activation of GM-CSF DCs.
GM-CSF DCs were stimulated as described for Figure 6.6. Expression of co-stimulatory molecules was determined by flow cytometry. The median fluorescence intensity (MFI) was calculated using FlowJo. Data shown are the mean from triplicate samples $\pm$ s.d. Statistical analysis using Student's *t*-test showed that there was no significant increase in expression of any of the molecules from medium only controls following stimulation with R:D60 (CD40: R:D60 $p = 0.566$, R:D60 + Lipofectamine $p = 0.137$; CD80: R:D60 $p = 0.312$, R:D60 + Lipofectamine $p = 0.884$; CD86: R:D60 $p = 0.643$, R:D60 + Lipofectamine $p = 0.469$; MHC Class II: R:D60 $p = 0.528$, R:D60 + Lipofectamine $p = 0.787$).

6.6. Stimulation of Flt-3L-DCs with R:D60

6.6.1. FL-DC bone marrow cultures

As discussed in Chapter 1, pDCs can produce large amounts of type I IFNs in response to both pathogenic and host nucleic acids (reviewed by Lande and Gilliet, 2010). *In vivo*, pDCs are a rare cell population, constituting less than 1% of peripheral blood cells. The number of both pDCs and cDCs in the blood can be greatly increased by the injection of Fms-like tyrosine kinase 3-ligand (Flt3-L) (Maraskovsky *et al.*, 1996; Pulendran *et al.*, 2000).

In vitro, murine pDCs can be generated from bone marrow cultures containing Flt3-L (FL-DC cultures) (Brawand *et al.*, 2002; Gilliet *et al.*, 2002). FL-DCs are a heterogeneous population of $CD11c^{+}$ cDCs and pDCs which can be segregated into distinct populations by their expression of surface markers (Figure 6.9). These are equivalent to steady-state splenic pDC and cDC populations, although the relative proportions of each population are slightly altered (Naik *et al.*, 2005). pDCs express B220 (CD45RA) but cDCs do not (Figure 6.9 **a**) (Brawand *et al.*, 2002; Gilliet *et al.*, 2002). The cDC population can further be subdivided into $CD24^{hi}$ and $CD11b^{hi}$/Sirp-α^{hi} populations, which correspond to $CD8^{+}$ and $CD8^{-}$ splenic cDC populations, respectively(Naik *et al.*, 2005) (Figure 6.9 **b**). The three different DC subsets present in FL-DC cultures exhibit distinct patterns of TLR expression and produce specific cytokines in response to stimulation with TLR ligands (Brawand *et al.*, 2002; Gilliet *et al.*, 2002; Naik *et al.*, 2005) (summarised in Table 6.1).

To investigate if R:D60 could stimulate a cytokine response in pDCs, FL-DC cultures were transfected in a repeat of the experiments with macrophages and GM-CSF cDCs. All R:D60 used in the following experiments had been FPLC-purified and verified by native PAGE analysis. As in the previous experiments in this chapter, the TLR9 ligand CpG ODN (A type), which is known to induce high production of IFN-α by pDCs without requiring transfection (Krug *et al.*, 2001), was included as a control.

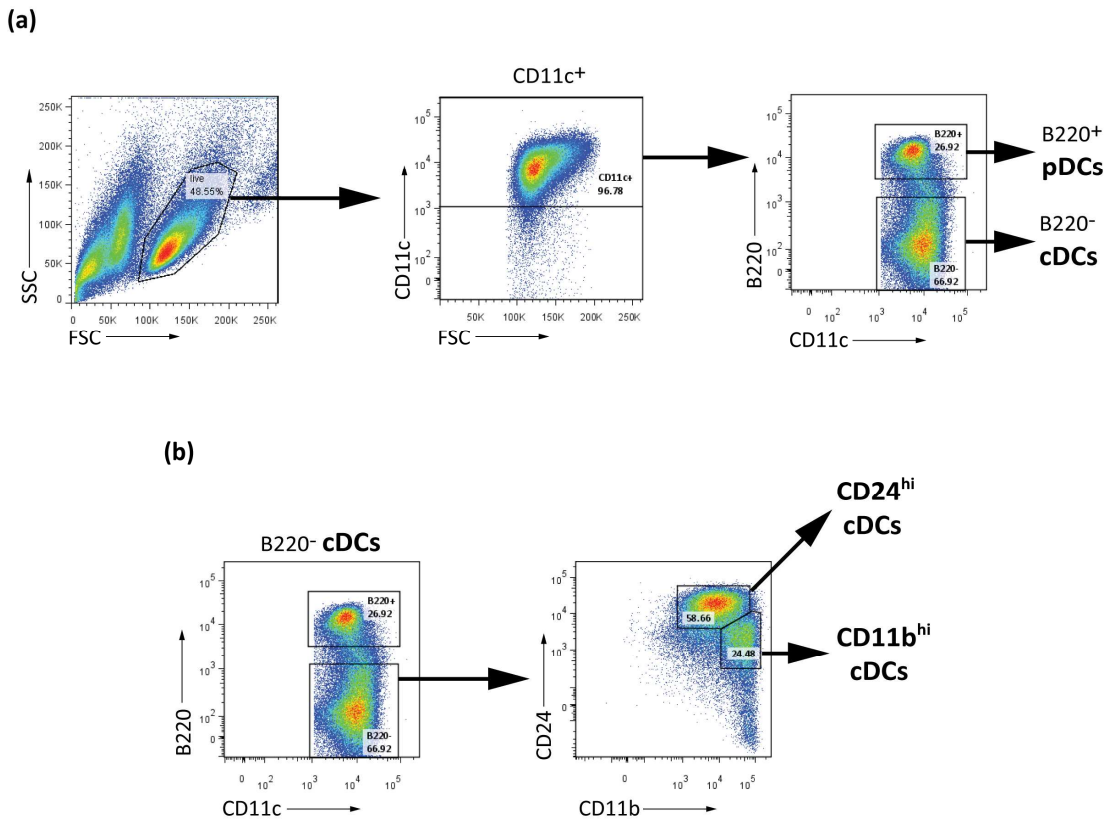


Figure 6.9. Segregation of FL-DCs into DC subsets using surface marker expression. **(a)** $CD11c^+$ cells were separated using B220/CD45RA staining into two populations: $B220^+$ (pDCs) and $B220^-$ (cDCs). During the experiments is this thesis, typically in the range of 20 – 50% of cells from each Flt3-L culture were $B220^+$. **(b)** The cDCs could be subdivided into $CD24^{hi}$ and $CD11b^{hi}$ populations. Data shown are representative of multiple experiments. The relative proportions of the populations did not alter upon stimulation with nucleic acids, in agreement with previously reported data (Brasel *et al.*, 2000).

Table 6.1. Comparison of the characteristics of the DC subpopulations present in Flt3-L bone marrow cultures.

	pDCs	cDCs (CD8+ equivalent)	cDCs (CD8- equivalent)
Surface markers	$CD11c^+$ $B220^+$	$CD11c^+$ $B220^-$ $CD24^{hi}$	$CD11c^+$ $B220^-$ $CD11b^{hi}$
CD8 mRNA expression	++	+	-
CD4 mRNA expression	++	-	-
TLR expression			
TLR3	-	+++	-
TLR4	-	++	+++
TLR7	+++	-	+
TLR9	+++	++	++
Chemokine receptor expression			
CXCR3	+++	++	+
CCR9	+++	-	-
CCR6	+	-	++
CXC3CR1	+	+	+++
Cytokine production			
R848	RANTES, MIP-1α	-	IL-12p40, RANTES
CpG ODN (Type A)	IFN-α	IL-12p40, IL-12p70	IL-12p40

Table assembled using data from Naik *et al*. and Brawand *et al*.

6.6.2. Transfected R:D60 stimulates cytokine production by FL-DCs

FL-DC cultures containing 20 – 50% $CD11c^+$ $B220^+$ pDCs were stimulated with the RNA:DNA hybrid R:D60, in the presence or absence of Lipofectamine. In contrast to macrophages and GM-CSF DCs, FL-DCs secreted cytokines in response to transfected R:D60 (Figure 6.10). In particular, high levels of IFN-α were detected in the supernatants of transfected cells. IFN-β was also produced in response to transfected R:D60, as were IL-6, RANTES and TNF-α. Low levels of the anti-inflammatory cytokine IL-10 were detected, but were not significantly increased from the Lipofectamine only control (Student's *t*-test, $p = 0.509$).

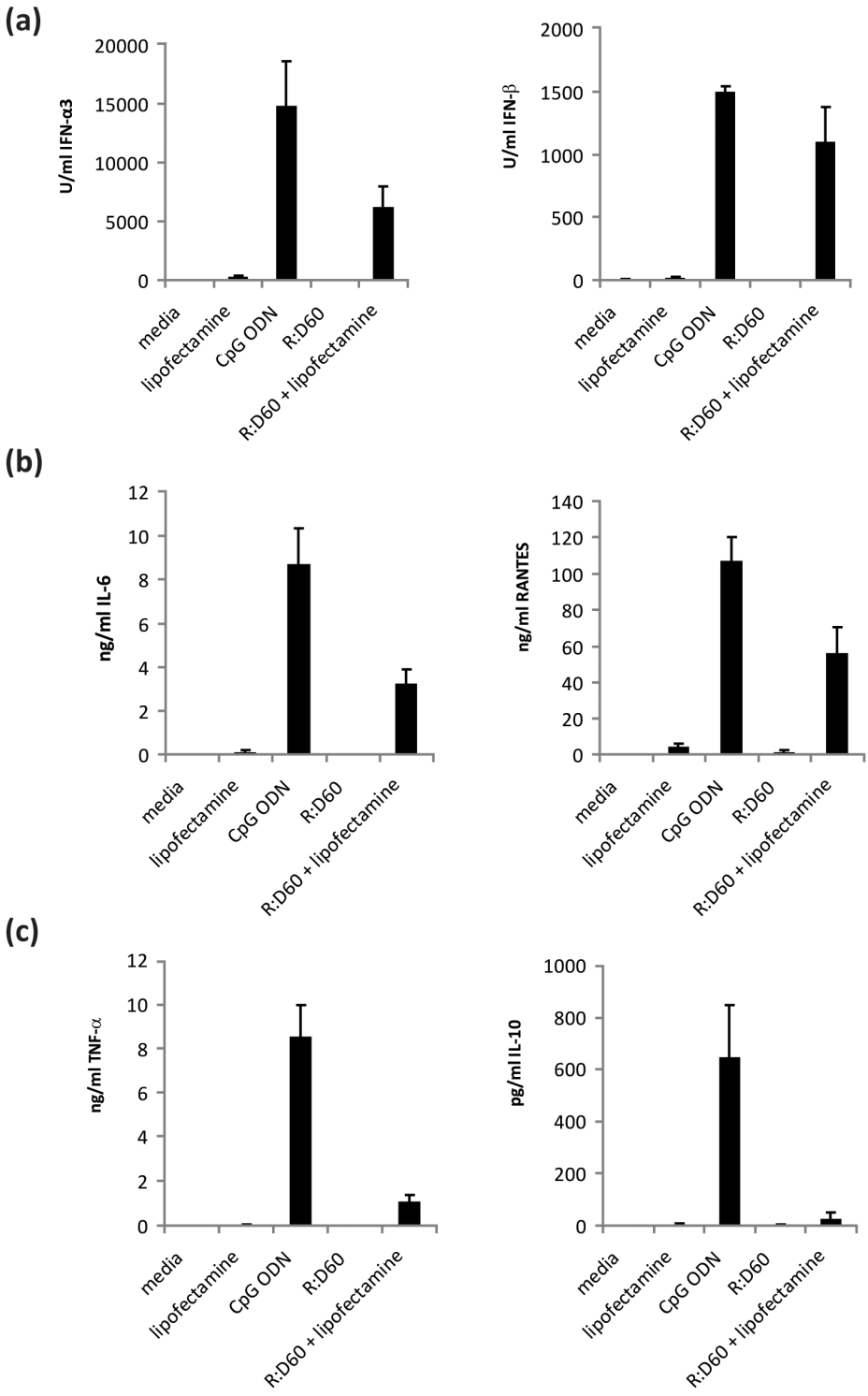


Figure 6.10. Transfected R:D60 stimulates cytokine production by FL-DCs.
FL-DC cultures were stimulated with 1 µg/ml CpG ODN, 1 µg/ml R:D60 or were transfected with 1 µg/ml R:D60 using Lipofectamine. After 18 h, supernatants were collected and secreted cytokines detected by ELISA. Transfected R:D60 stimulates the production of type I IFN **(a),** IL-6 and RANTES **(b)** by FL-DCs. Some TNF-α and IL-10 are also produced **(c)**. Data shown are the mean of five (IFN-α and IL-6), four (IFN-β and TNF-α), three (RANTES) or two (IL-10) independent experiments ± s.e.m.

6.6.3. FL-DCs become phenotypically activated following transfection with R:D60

Unstimulated cells exhibited an immature phenotype, with low expression of CD40, CD80, CD86 and MHC Class II; however, expression of all four was upregulated by cDCs and of all except CD80 by pDCs in response to transfected R:D60 (Figure 6.11). These cells also upregulated expression of CD8, as previously reported for FL-DC cultures activated with TLR ligands (Brasel *et al.*, 2000; Brawand *et al.*, 2002).

In the absence of Lipofectamine, R:D60 failed to induce cytokine production by FL-DC cultures (Figure 6.10) or expression of activation markers by pDCs, although a low level of phenotypic activation was observed in cDCs (Figure 6.11). Thererfore, internalisation of R:D60 is required for cytokine production by FL-DCs.

6.6.4. Type I IFN genes are upregulated in FL-DCs transfected with R:D60

A preliminary analysis of cytokine gene expression in FL-DCs by qRT-PCR showed that transfected R:D60 strongly induces type I IFN gene expression (Figure 6.12), the profile of which differed from that induced by CpG ODN. The fold induction of *Ifna1* mRNA had fallen by 18 h in response to the hybrid, however it was sustained following CpG ODN stimulation. Similarly, expression of *Ifnb1* increased over time following stimulation with CpG ODN but did not in response to the hybrid. Fold induction of pro-inflammatory cytokine gene expression was much lower, as shown for *Tnfa*. No increase in *Il1b* expression was detected in response to R:D60 (Figure 6.12). Together with the ELISA data (Figure 6.10), this suggested that the cytokine response to transfected R:D60 by FL-DC cultures is substantially a type I IFN response.

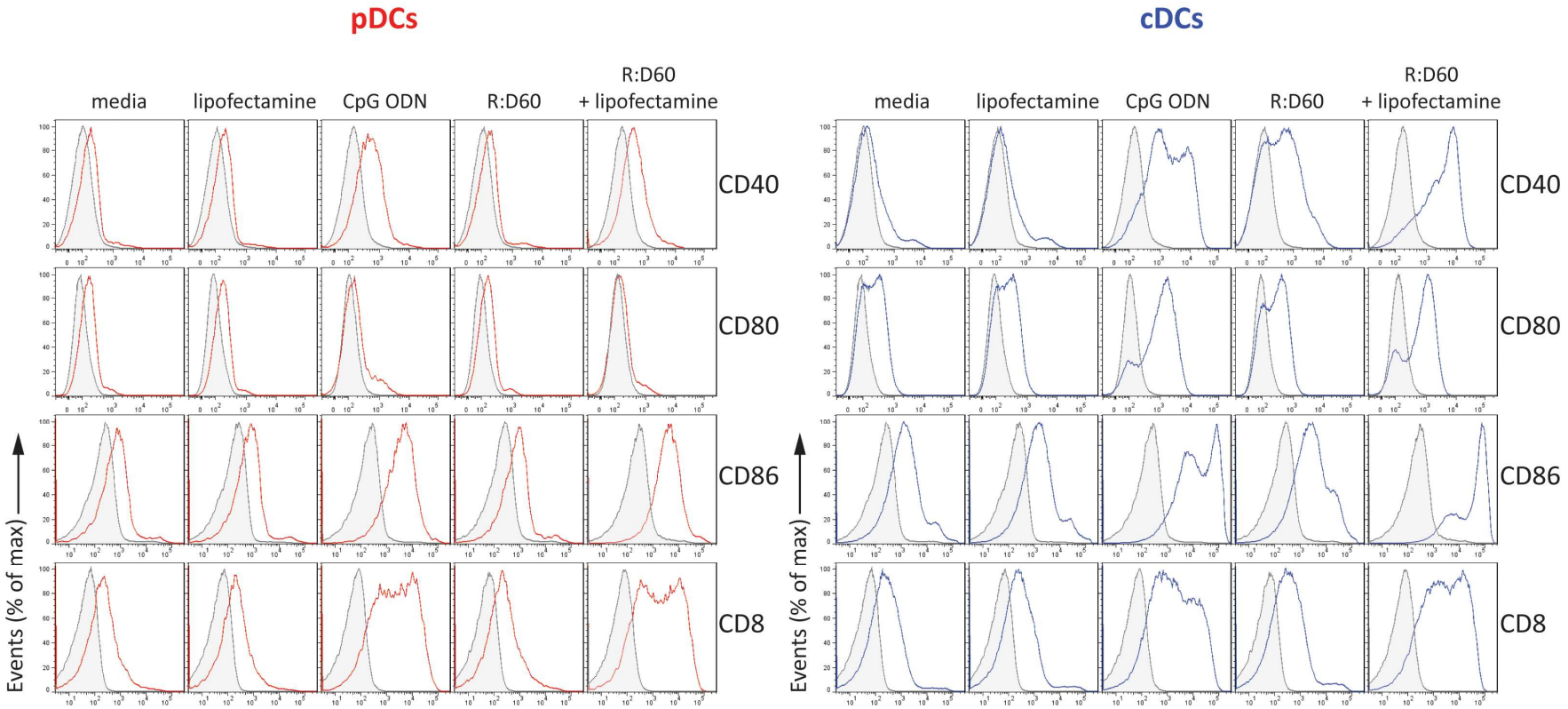


Figure 6.11. R:D60 stimulation activates FL-DCs.
FL-DC cultures were stimulated with 1 µg/ml CpG ODN, 1 µg/ml R:D60 added to the medium or were transfected with 1 µg/ml R:D60 using Lipofectamine 2000. After 18 h, cell phenotype was assessed by flow cytometry. Transfected R:D60 induced an upregulation of the costimulatory molecules CD40 and CD86 in both pDCs and cDCs, as did stimulation with CpG ODN. Cell surface expression of CD8 was also upregulated by CpG ODN and transfected R:D60. Isotype controls are shown in grey. Data are representative of four independent experiments.

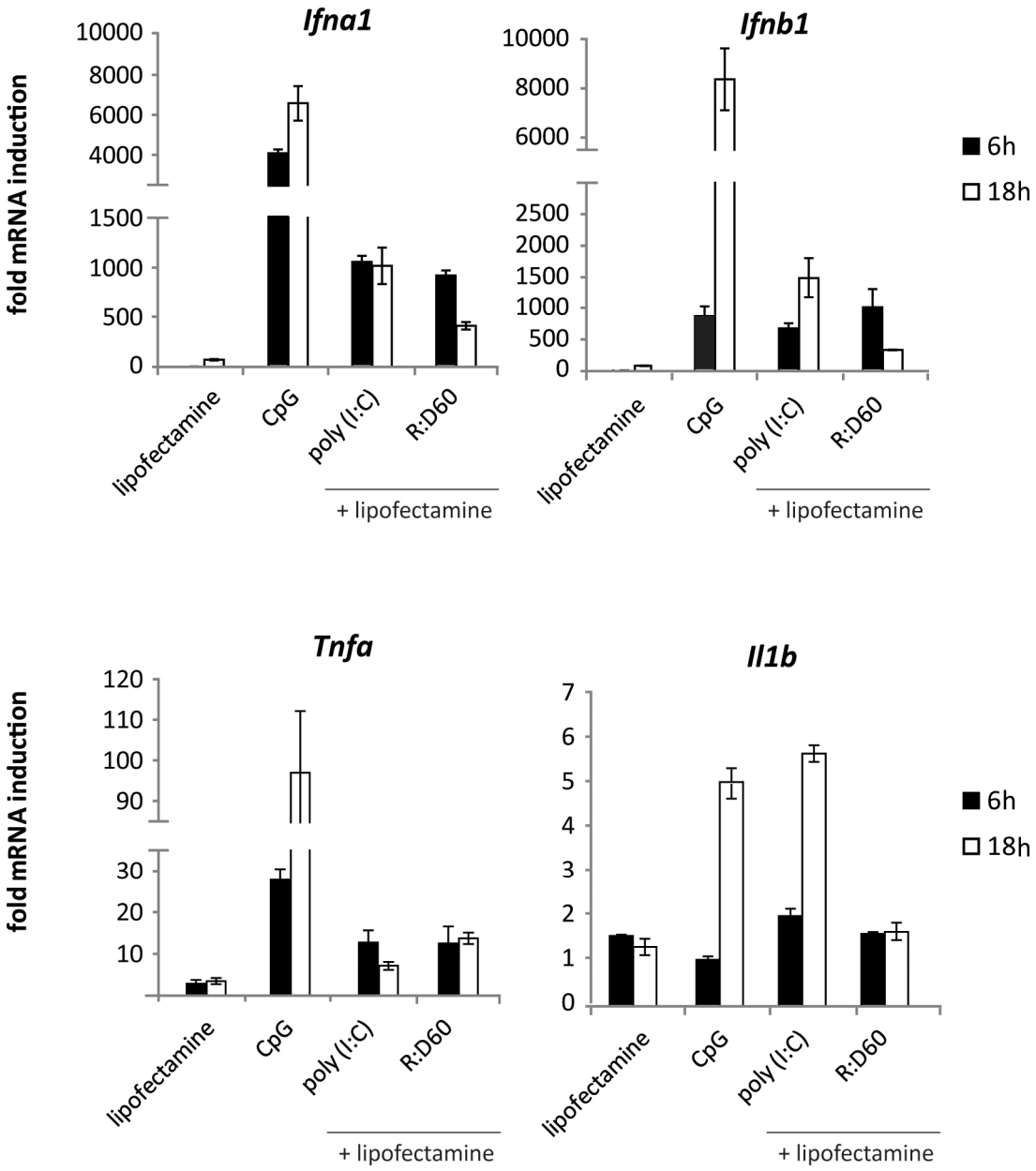


Figure 6.12. Transfection of R:D60 into FL-DCs strongly induces expression of type I IFN genes.

FL-DCs were stimulated with 1 µg/ml CpG ODN or transfected with 1 µg/ml poly (I:C) or 1 µg/ml R:D60 using Lipofectamine 2000 for the indicated times before qRT-PCR analysis was performed. Data are from one experiment, represented as fold mRNA increase from the medium only control ± s.d. of triplicate qPCR wells.

6.6.5. Optimisation of FL-DC transfection experiments

R:D60 stimulated type I IFN production in FL-DC cultures when introduced into cells using cationic-lipid mediated transfection. It was important to establish that the response is due to the hybrid rather than any contamination introduced into the cells

6.6.5.1. Lipofectamine 2000 can activate FL-DCs

It has previously been reported that Lipofectamine 2000 does not induce non-specific cytokine production when used to transfect nucleic acids into immune cells, specifically DCs (Diebold *et al.*, 2003). Initial experiments using Lipofectamine 2000 were found to agree with this, however over the course of multiple repeats of these transfection experiments, it was observed that by itself Lipofectamine 2000 could sometimes cause low levels of cytokine production (Figure 6.13 **a**) and induce phenotypic activation of the cells (Figure 6.13 **b** and **c**). The extent of non-specific activation of FL-DCs was found to be variable dependent on the batch of Lipofectamine 2000. The stimulation of type I IFN production by transfection reagents has been reported previously (Li *et al.*, 1998). However, the use of a variant of Lipofectamine (Lipofectamine LTX), which the manufacturers recommend for the transfection of sensitive cell lines, minimised this non-specific activation, and confirmed that transfected R:D60 did specifically induce a response in these cells (Figure 6.13). Therefore, Lipofectmine LTX provided the better option for the transfection of FL-DCs.

6.6.5.2. Dose response test

A dose response experiment showed that the levels of cytokines produced and surface marker expression are similar when 0.5 – 5 μg/ml of R:D60 is transfected into FL-DC cultures (Figure 6.14), although the highest dose did stimulate more IFN-α production in this experiment. Therefore, to reduce the level of any residual contamination of ssNA in the hybrid preparation being introduced into the cell, 1μg/ml of hybrid was continued to be used for all subsequent experiments.

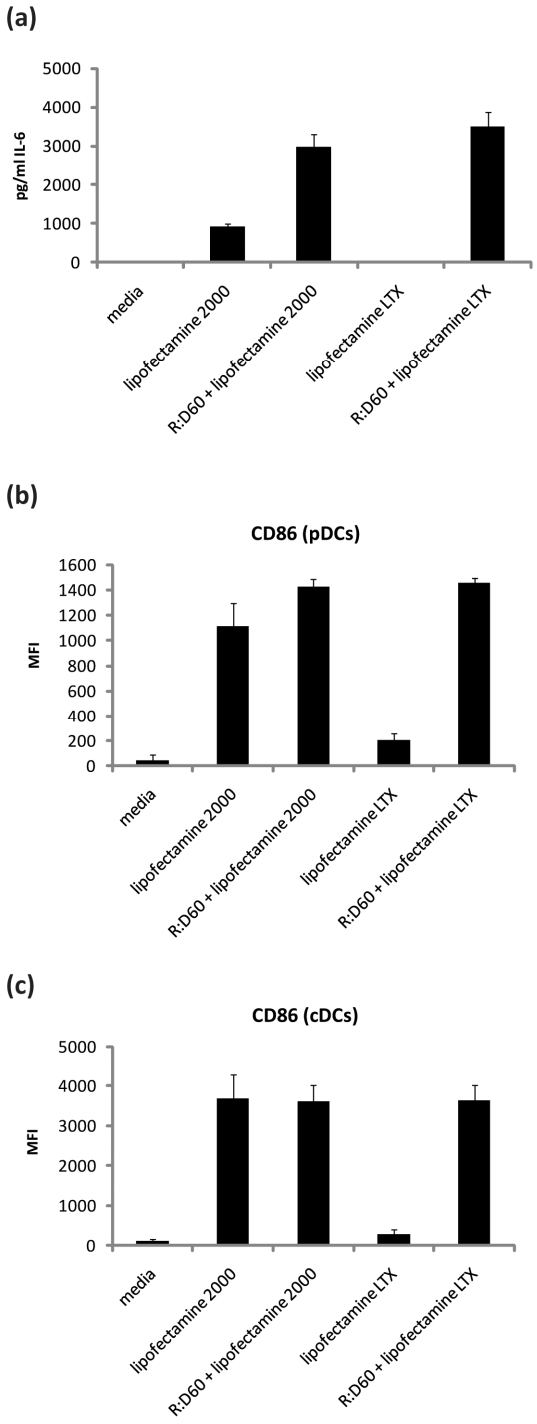


Figure 6.13. Comparison of non-specific activation of FL-DCs by Lipofectamine 2000 and Lipofectamine LTX.

FL-DCs were transfected with 1 μg/ml R:D60 using Lipofectamine 2000 and Lipofectamine LTX. Lipofectamine only controls were also included. After 18 h, levels of cytokine production were determined by ELISA and expression of activation markers by flow cytometry. Lipofectamine 2000 alone treatment, but not Lipofectamine LTX, stimulated the production of cytokine, as shown for IL-6 **(a)**. Expression of activation markers in both pDCs and cDCs were also increased in these cells, as shown for CD86 **(b)** and **(c)**. Data shown are from duplicate samples ± s.d.

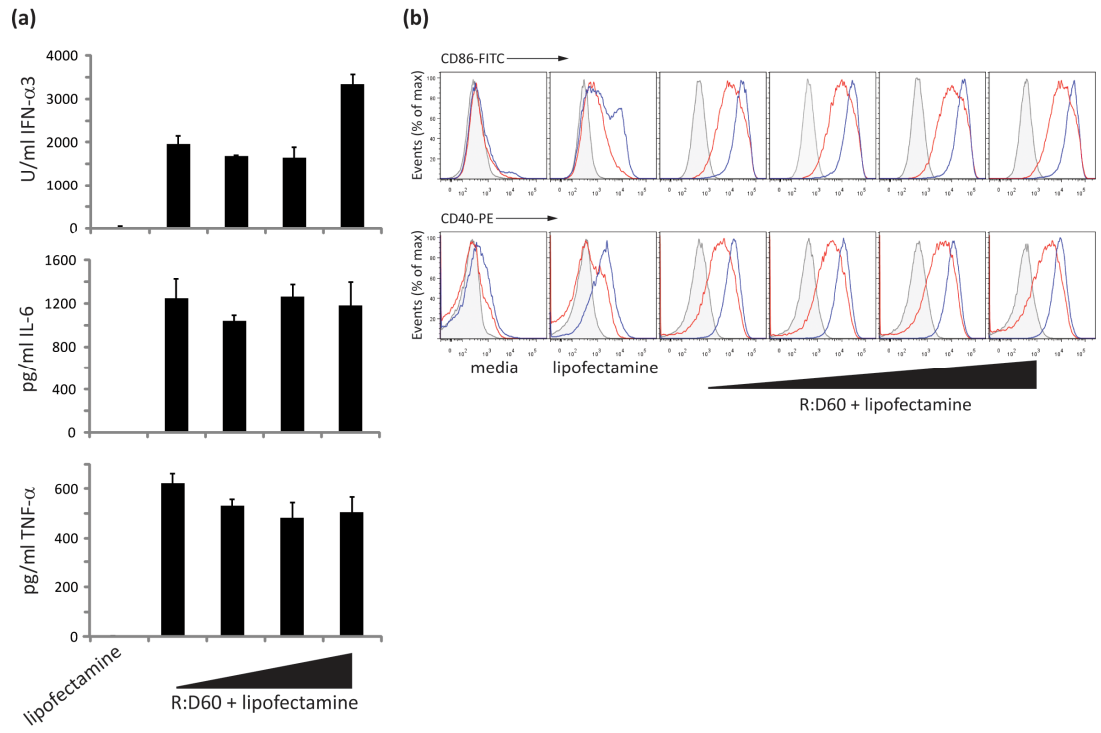


Figure 6.14. The response of FL-DCs to different concentrations of R:D60.
FL-DCs were transfected with increasing concentrations of R:D60 (0.5, 1, 2 and 5 µg/ml) using Lipofectamine 2000. After 18 h, supernatants were collected and the cells stained for flow cytometry. **(a)** Levels of cytokines produced by FL-DCs are the same as for 0.5 – 2 µg/ml, but more IFN-α is produced in response to 5 µg/ml R:D60. Data shown is the mean of triplicate samples ± s.d. **(b)** The upregulation of co-stimulatory molecule expression is equivalent for all concentrations of R:D60 used, as shown for CD40 and CD86. Isotype controls are shown in grey, $CD11c^+$ $B220^+$ (pDCs) in red and $CD11c^+$ $B220^-$ (cDCs) in blue. Data shown is representative of triplicate samples.

The data in Figure 6.14, although preliminary, suggests that there is no clear effect on levels of cytokine production and phenotypic activation with increasing doses of the hybrid. An explanation for this could be the establishment of type I IFN feedback loops, where early detection of the hybrid in individual cells triggers IFN production, which then stimulates further type I IFN production via engagement of the type I IFN receptor (IFNAR). This would amplify the IFN-induced stimulation of other cytokines in cells which are directly responding to the hybrid and, in addition, trigger type I IFN production by other cells. Therefore, this autocrine-paracrine loop would generate large amount of cytokines, even by cells that have not been transfected. Therefore, the transfection of increased concentrations of R:D60 may not stimulate a

noticeable difference in cytokine production and phenotypic activation from cells transfected with low amounts. However, the use of 5 μg/ml R:D60 did stimulate more IFN-α production (Figure 6.14 **a**).

6.6.6. Validation of the response of R:D60

The stimulation of FL-DCs by transfected R:D60 represented the first time an RNA:DNA hybrid has been shown to activate immunity, therefore it is not known which receptor pathway would be involved. The production of type I IFNs is characteristic of an innate immune response to nucleic acids, particularly by pDCs. As detailed in Chapter 5, the production of R:D60 involves the annealing of ssDNA and ssRNA oligonucleotides. Purification of the hybrid using FPLC and subsequent analysis by native PAGE should reduce/prevent contamination of the R:D60 preparations by unannealed ssNA. If not removed, it is conceivable that the ssDNA and ssRNA oligonucleotides may stimulate type I IFN production in pDCs, possibly via TLR7 and TLR9. To test the immunostimulatory potential of the ss oligonucleotides, FL-DCs were transfected with different amounts of oligonucleotides and IFN-α measured by ELISA (Figure 6.15).

When transfected at a concentration of 500 ng/ml, ssRNA60 strongly stimulated IFN-α production by FL-DCs (Figure 6.15 **a**), potentially via TLR7 since it is GU-rich ssRNA. In contrast, ssDNA60 failed to induce high levels of IFN-α when used at this concentration. 1 μg of R:D60 would theoretically contain a maximum of 500 ng ssDNA60. Therefore IFN-α production in response to 1 μg R:D60 is unlikely to be induced by ssDNA60 contamination. When used at 1 μg/ml, both ssRNA60 and ssDNA60 stimulated IFN-α production in FL-DCs at levels at least equivalent to those following R:D60 stimulation (Figure 6.15 **b**).

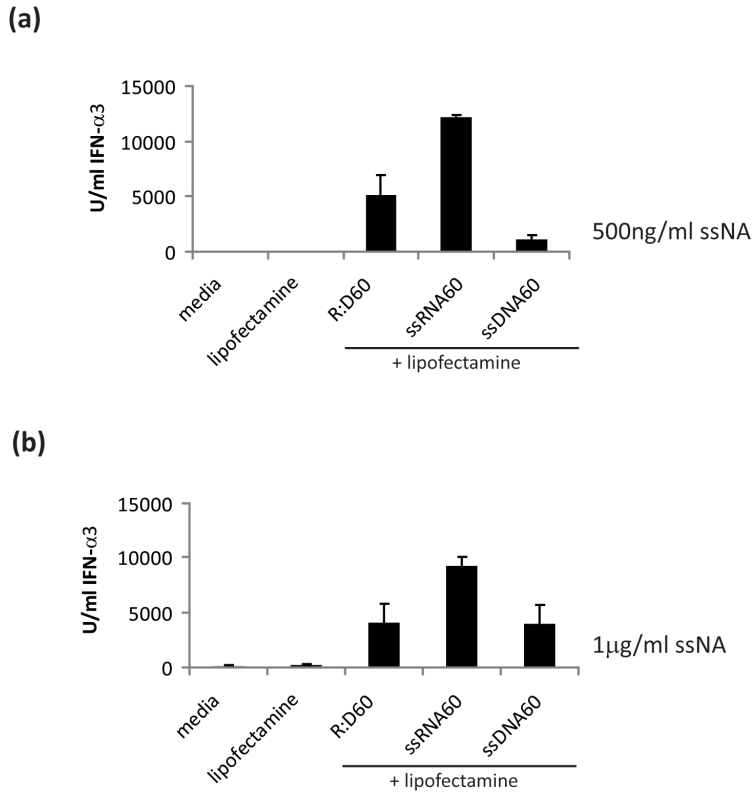


Figure 6.15. Stimulation of IFN-α production by FL-DCs with different concentrations of single-stranded nucleic acids.
FL-DC cultures were transfected with 1 μg/ml R:D60 and different concentrations of ssRNA60 and ssDNA60. **(a)** and **(b)** are each from independent experiments performed in duplicate, error bars are ± s.d.

A preliminary experiment suggested that the cytokine response to R:D60 is not abolished when the hybrid is digested with *E. coli* RNase H1 (Figure 6.16.). RNase H1 hydrolyses the RNA strand of RNA:DNA hybrids in a distributive manner, periodically cleaving the RNA to leave small fragments. Digestion of 1 μg of R:D60 by RNase H1 therefore resulted in polyribonucleotide fragments, in addition to the 500 ng of undigested DNA strand. As shown in Figure 6.15, ssRNA60 was highly immunostimulatory, even at 500 ng/ml. Therefore, it could be that the short RNA products of RNase H digestion are responsible for the stimulation of IFN-α production in this experiment.

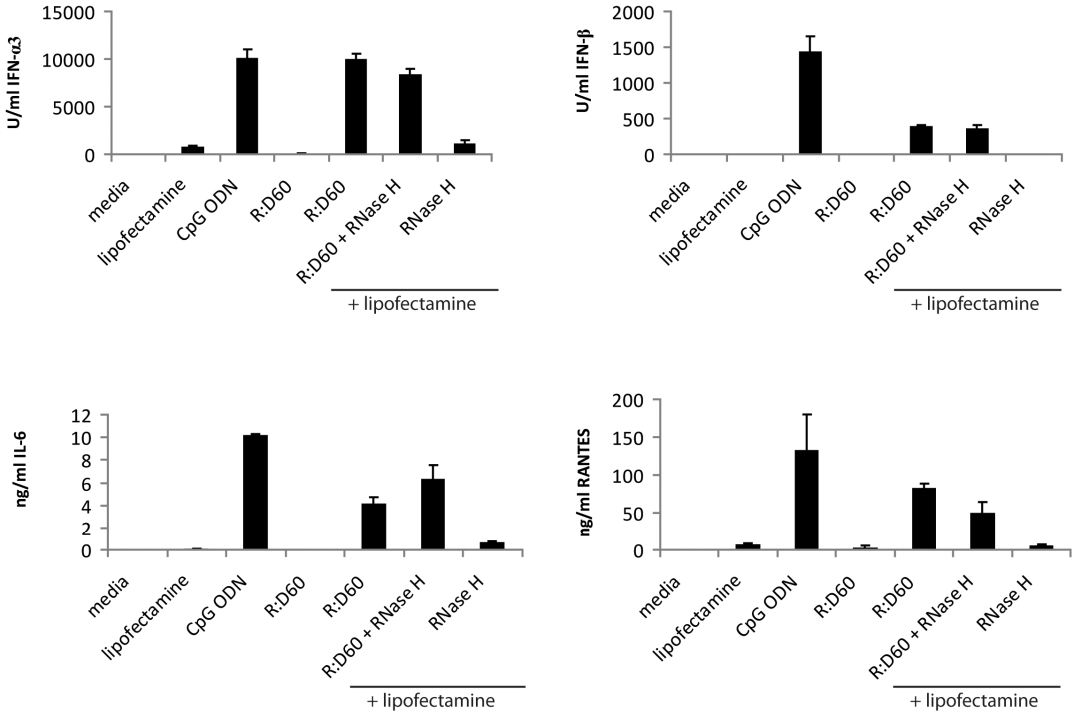


Figure 6.16. Digestion of R:D60 with RNase H does not abrogate the cytokine response by FL-DCs.

R:D60 was digested with RNase H as previously (see Figure 5.7). R:D60 minus RNase H and RNase H alone were included as controls. Digests were ethanol precipitated and resuspended to equivalent concentrations. FL-DCs were stimulated with 1 µg/ml CpG ODN or 1 µg/ml R:D60 by addition to the culture medium. 1 µg/ml undigested R:D60 and an equal volume of R:D60 + RNase H and the enzyme only control were transfected into cells using Lipofectamine 2000. After 18 h, supernatants were assayed by ELISA. Data shown are from one experiment ± s.d. of triplicate ELISA wells.

Finally, the possibility that immunostimulatory contaminants were introduced during the FPLC process was examined. Oligonucleotide annealings containing either both oligos (R:D60), only the DNA oligo (sham RNA) or only the RNA oligo (sham DNA) were prepared. Native PAGE analysis confirmed that the sham annealings contained only one strand (Figure 6.17 **a**). Each annealing mix was fractionated using FPLC (Figure 6.17 **b**). The sham mixes, which contained only ssNA, eluted later than R:D60 (Figure 6.17 **b** and Figure 5.5). The fractions where the hybrid would elute were collected ('R:D fractions') as well as the peaks. Following ethanol precipitation, all fractions were resuspended in the same volume of nuclease-free water and analysed by native PAGE (Figure 6.17 **c**). In both sham annealings, no

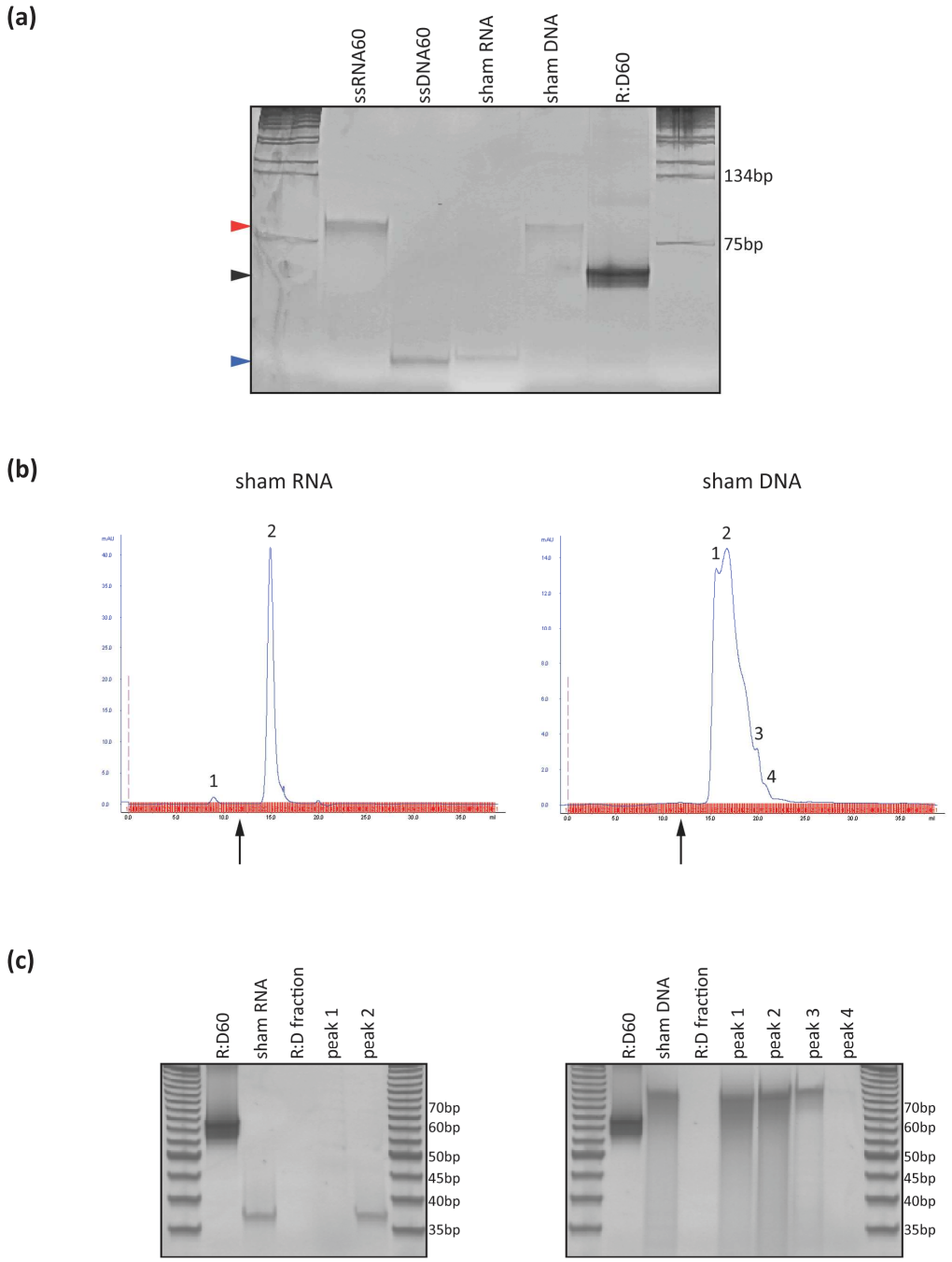


Figure 6.17. FPLC controls for FL-DC transfections.
(a) Native PAGE analysis of R:D60 and sham annealings. 500 ng each of ssDNA60, ssRNA60, R:D60 and the equivalent volume of sham RNA (DNA oligonucleotide only) and sham DNA (RNA oligonucleotide only) were electrophoresed on a 12% polyacrylamide gel, which was then stained with acridine orange. The arrowheads indicate the migration of ssRNA60 (red), ssDNA60 (blue) and R:D60 (black). Note faint white line co-migrating with ssDNA60 is the reference dye front. **(b)** FPLC analysis of the sham annealings. Black arrows indicate the point where the hybrid elutes (R:D fractions) (see Figure 5.4). **(c)** Native PAGE analysis of sham annealing FPLC fractions. 200 ng of FPLC purified R:D60, 200 ng of sham RNA or sham DNA input and equivalent volumes of ethanol precipitated fractions were electrophoresed on a 15% native polyacrylamide gel. Nucleic acids were visualised by Sybr Gold staining.

nucleic acid was detected in the R:D fraction by PAGE analysis (Figure 6.17 **c**) which was verified by quantification using nucleic acid absorbance at 260 nm.

FL-DCs were transfected with 1 μg/ml R:D60 and equivalent volumes of the R:D fractions from both of the sham annealings. The R:D fractions from the sham annealings failed to stimulate cytokine production in FL-DCs (Figure 6.18 **a** and **b**). In addition, they failed to induce upregulation of co-stimulatory molecule expression above Lipofectamine only controls in both pDCs and cDCs, as shown for CD86 in Figure 6.18 **a** and **b** (it should be noted that in these experiments Lipofectamine 2000 alone did stimulate upregulation, which complicates the interpretation of data). Therefore, the cytokine response of FL-DCs to transfected R:D60 is unlikely to be due to immunostimulatory contaminants introduced by the FPLC purification process.

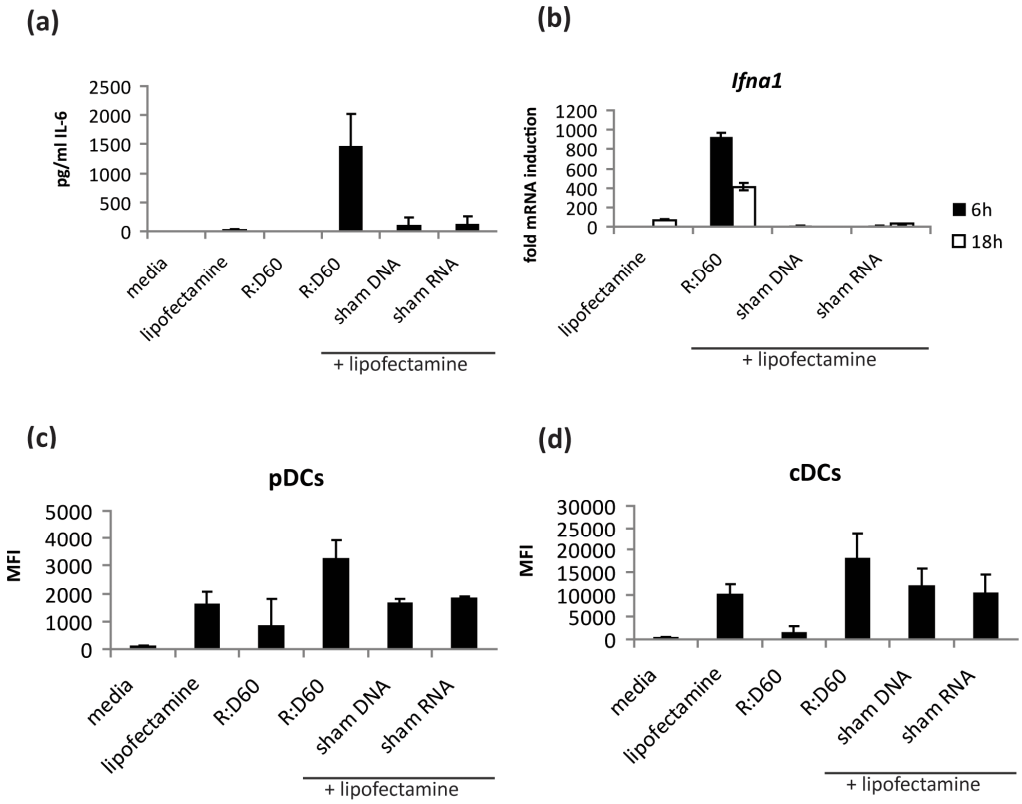


Figure 6.18. The FPLC R:D fractions from sham annealings do not stimulate FL-DCs. FL-DCs were transfected with 1 µg/ml R:D60 and the equivalent volume of R:D fractions from FPLC of sham annealings, using Lipofectamine 2000. **(a)** Sham annealing R:D fractions did not stimulate IL-6 production above background (Lipofectamine only controls) in FL-DC cultures (Student's *t*-test; R:D60 $p = 0.054$, sham RNA $p = 0.481$, sham DNA $p = 0.489$). Data shown are the mean of three independent experiments ± s.e.m. **(b)** Sham annealing R:D fractions did not induce expression of *Ifna1* in FL-DCs. Data are from one experiment ± s.d. of qRT-PCR triplicates. **(c)** and **(d)** Sham annealing R:D fractions did not increase expression of CD86 from Lipofectamine 2000 only controls. MFIs were calculated using FlowJo. Data shown are the mean of two independent experiments ± s.d.

6.6.7. The stimulation of innate immunity in response to other RNA:DNA hybrids

To investigate if the response of FL-DCs to R:D60 can be replicated using other RNA:DNA hybrids, preliminary experiments using poly (rA:dT) and R:D20 were performed.

6.6.7.1. Poly (rA:dT) does not activate FL-DCs

The commercially available RNA:DNA hybrid poly (rA:dT) is frequently used in RNA:DNA hybrid structural studies and is a known substrate for RNase H enzymes *in vitro* (Shaw and Arya, 2008). Poly (rA:dT) is generated by annealing equimolar amounts of each polymer. The individual polymers are synthesised using enzymes, so are polydisperse. Consequently, there is extensive variability in size between different commercially available batches of poly (rA:dT) but typically all contain hybrids of at least a few hundred bp in length (Richard V Case, Midland Certified Reagent Company, Inc., personal communication).

Lyophilised poly (rA:dT) was reconstituted in nuclease-free water and used to stimulate FL-DCs in the presence and absence of Lipofectamine. Transfected poly (rA:dT) failed to activate FL-DCs above background levels (Lipofectamine only controls) (Figure 6.19 **a**), and did not induce a cytokine response (Figure 6.19 **b**).

6.6.7.2. R:D20 stimulates cytokine production in FL-DCs

To determine if shortening the length of R:D60 affected its ability to stimulate a cytokine response in FL-DCs, a preliminary experiment using an FPLC-purified 20 bp hybrid (R:D20) consisting of the same sequence and nucleic acid chemistry as R:D60 was performed. This hybrid stimulated type I IFN production at least as well as R:D60 (Figure 6.20). However, although R:D20 definitely contains RNA:DNA hybrid, PAGE analysis performed in Chapter 5 suggest that it could contain more than one nucleic acid species (Figure 5.6). Nevertheless, these data potentially suggest that recognition of $r(GU)_n$:$d(CA)_n$ duplexes could be dependent on the sequence or conformation of the hybrid rather than the length.

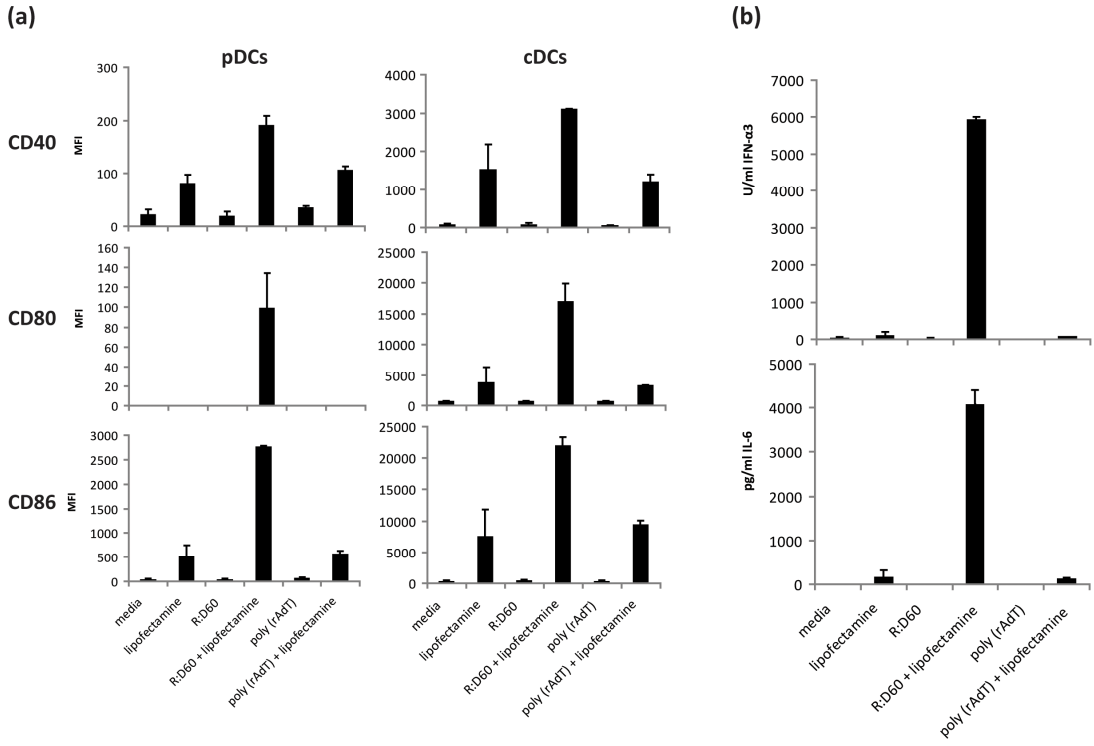


Figure 6.19. FL-DCs do not respond to poly (rA:dT).
FL-DCs were stimulated with 1 µg/ml R:D60 and 1 µg/ml poly (rA:dT) ± Lipofectamine 2000 for 18 h. **(a)** Expression of CD40, CD80 and CD86 was measured by flow cytometry and the MFI calculated for each sample. Note high background levels of expression induced by Lipofectamine 2000 alone. **(b)** IFN-α and IL-6 production were measured by ELISA. Data shown are from one experiment performed in duplicate ± s.d.

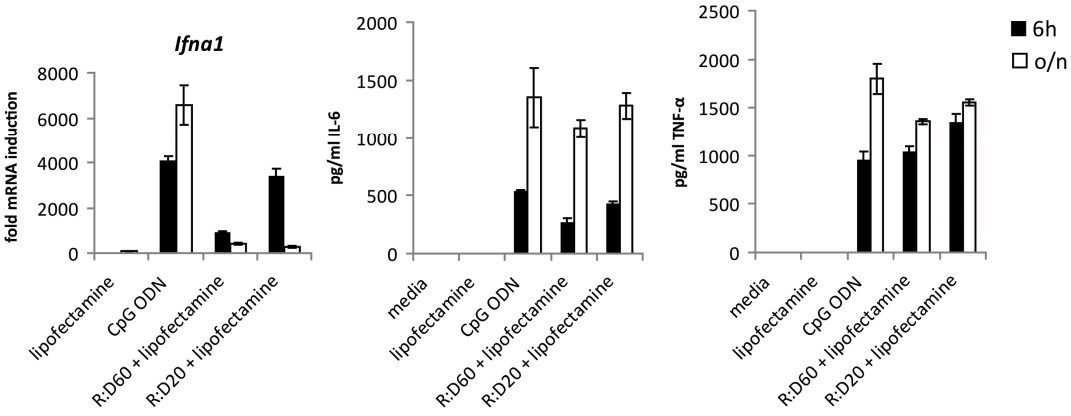


Figure 6.20. Cytokine production is stimulated by a 20bp RNA:DNA hybrid.
FL-DCs were stimulated with 1 µg/ml CpG ODN or transfected with 1 µg/ml each of R:D60 and R:D20 using Lipofectamine LTX. After the indicated time points, RNA was extracted from the cells and supernatants harvested (o/n indicates an 18 h stimulation). *Ifna1* expression was determined by qRT-PCR and is represented as normalised fold increase in expression from medium only controls. IL-6 and TNF-α production were determined by ELISA analysis of the supernatants. All data are from one experiment ± s.d. of triplicate wells.

Taken together, the experiments using poly (rA:dT) and R:D20 imply that the FL-DC response to RNA:DNA hybrids could be sequence-dependent. However, the helical conformation of poly (rA:dT) is polymorphic; it is able to undergo A-form to B-form changes depending on the relative humidity of the environment (Shaw and Arya, 2008). As it is only an assumption that R:D20/R:D60 adopt an A-form conformation, the lack of response to poly (rA:dT) does not exclude the possibility that nucleic acid structure is important in innate immune recognition of RNA:DNA hybrids. The identification of the receptor responsible for intracellular detection of R:D60 and R:D20 would be advantageous in further refining RNA:DNA hybrid ligands.

6.6.8. Determination of the innate immune signalling pathways activated by R:D60 in FL-DCs

The two subtypes of DC present in FL-DC cultures are known to utilise different receptors in response to RNA viruses; pDCs use TLR7 and TLR9 signalling through MyD88, whereas cDCs rely on the RNA helicase-IPS-1 pathway (Kato *et al.*, 2005). Analysis of surface molecule expression by flow cytometry suggested that both subtypes are activated in response to transfected R:D60. However since FL-DCs are a heterogeneous population, it was not clear from the experiments performed whether both pDCs and cDCs can sense RNA:DNA hybrids or if one population becomes activated as a secondary effect.

Stimulation of a cytokine response by FL-DCs required transfection of R:D60 into the cell (Figure 6.10). As demonstrated in Figure 6.1, introduction of an 18 bp RNA:DNA hybrid into L929 cells using Lipofectamine localised the hybrid to the cytosol, although the precise localisation has yet to be determined. The TLR9 ligand CpG ODN does not require transfection to stimulate a cytokine response in pDCs; it is endocytosed and retained in the early endosomal compartments, where it is recognised by TLR9 (as discussed in Section 1.5.5.1). The failure of exogenous R:D60 to stimulate cytokine production suggested that the hybrid is unlikely to be endocytosed in a similar manner. However, pDCs can also respond to cytosolic nucleic acids, in a TLR-dependent manner. For instance, the transportation of viral replication intermediates from the cytosol to TLR7 in endosomes via autophagy provides a mechanism for pDCs to detect cytosolic RNA without utilising cytosolic receptors (Lee *et al.*, 2007). Therefore, to establish if the cytosolic RLRs were involved in the detection of R:D60 by FL-DCs, cultures from mice deficient in the mitochondrial membrane bound adaptor protein IPS-1 were stimulated with R:D60.

6.6.8.1. The response to R:D60 is independent of IPS-1

FL-DC cultures from IPS-$1^{-/-}$ mice and wild type controls on the same genetic background were stimulated with R:D60. The cytokine response of cells from the

IPS-1$^{-/-}$ mice was comparable with that of wild type controls (Figure 6.21), as was the extent of phenotypic activation (Figure 6.22). There was no detectable difference in the response of IPS-1$^{-/-}$ FL-DCs to the known RIG-I ligand IVT RNA, which presumably can be detected by TLR7 in pDCs. However, there was a reduction in the phenotypic activation of these cells in response to transfected poly (I:C), a Mda5 ligand (Appendix 11).

These data show that type I IFN production by FL-DCs in response to R:D60 is independent of IPS-1, and therefore the RNA helicases are unlikely to be the sensors for R:D60.

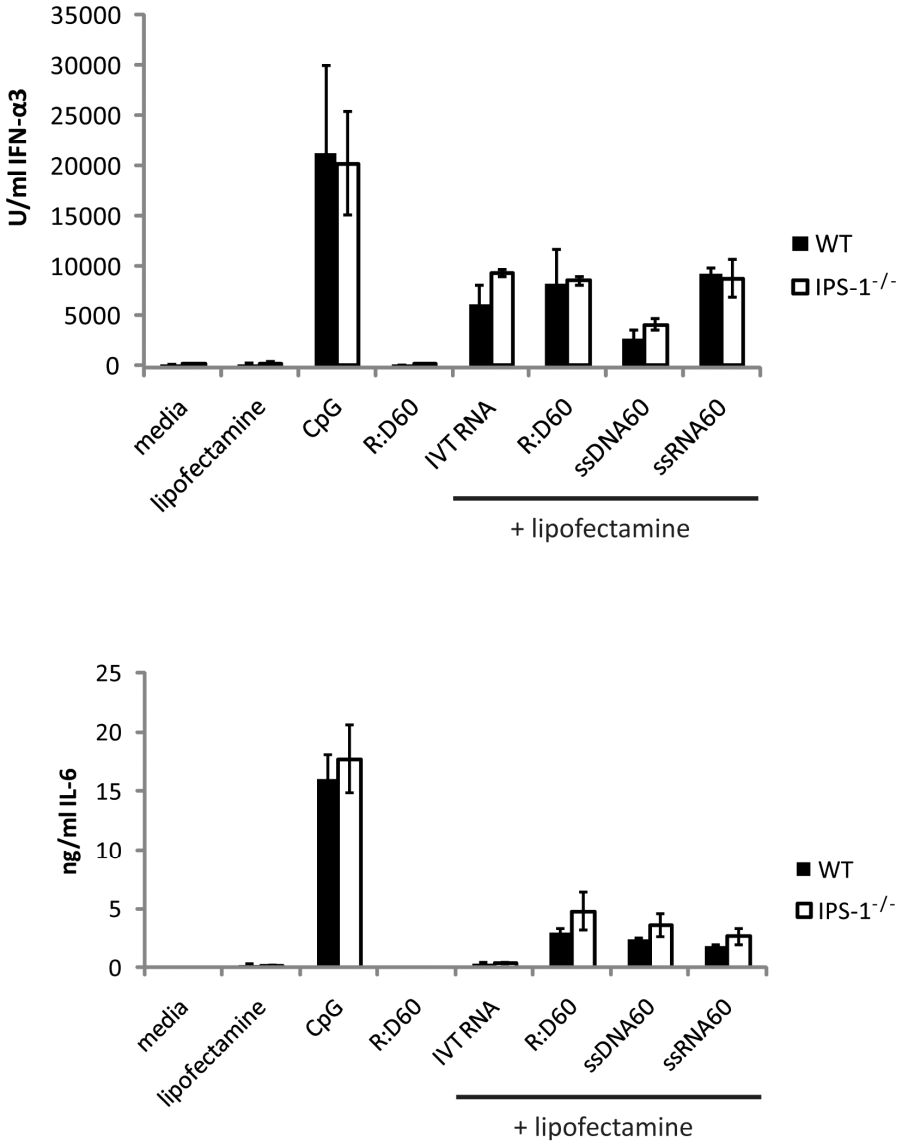


Figure 6.21. IPS-1 is not required for the cytokine response to transfected R:D60.
FL-DCs from IPS-$1^{-/-}$ mice and age/sex-matched wild type mice were stimulated with 1 μg/ml CpG ODN or 1 μg/ml R:D60, or were transfected with 1 μg/ml each of IVT RNA, R:D60, ssDNA60 or ssRNA60 using Lipofectamine 2000. Several nucleic acid ligands were included as controls, however IPS-$1^{-/-}$ FL-DCs produced equivalent levels of cytokines to wild type. There was no significant difference in the amount of IFN-α produced by wild type (WT) and IPS-$1^{-/-}$ cells in response to transfected R:D60 (Student's *t*-test, $p = 0.938$) or IL-6 ($p = 0.336$). Data are from three independent experiments ± s.e.m.

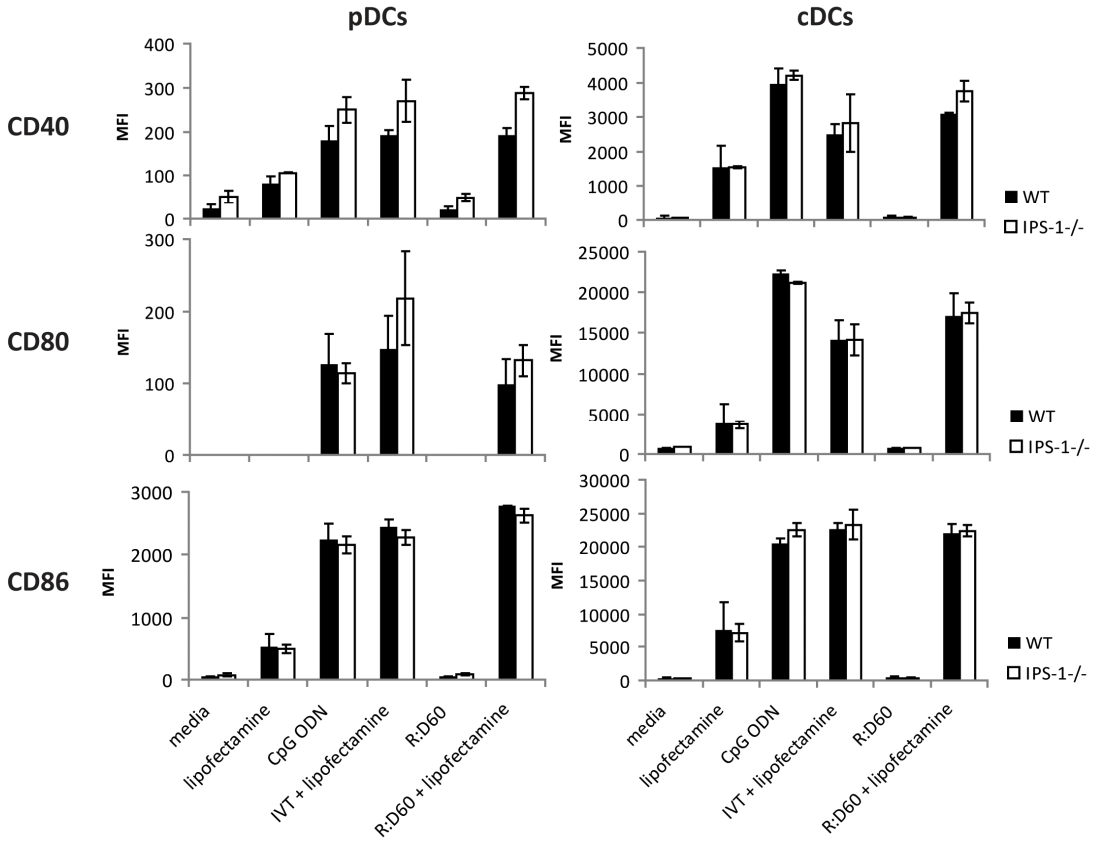


Figure 6.22. IPS-1 is not required for the phenotypic activation of FL-DCs by R:D60.
FL-DCs were stimulated as described for Figure 6.18. MFI values for CD40, CD80 and CD86 were calculated. Data shown are from two independent experiments ± s.d. Several nucleic acid ligands were included as controls, however there was no difference in expression between IPS-$1^{-/-}$ and WT cells stimulated with CpG ODN and IVT. There was a slight reduction in expression of CD86 in cells stimulated with the TLR3/MDA5 ligand poly (I:C) (Appendix 11).

6.6.8.2. Trif and/or MyD88 is required for the response of FL-DCs to R:D60

The intact immune response to R:D60 in cells lacking IPS-1 implicated a role for alternative pathways in the recognition and response of FL-DCs to R:D60, most notably TLR-dependent pathways. The nucleic acid-sensing TLRs require the adaptor molecules Trif (TLR3) or MyD88 (TLR7 and TLR9). pDCs do not express TLR3, however $CD24^{hi}$ cDCs do (Table 6.1). In addition, cytokine induction by the

two DEAH/RHA helicases DHX9 and DHX36 also requires MyD88 (Kim *et al.*, 2010). Therefore, to establish if the nucleic acid TLRs and/or these cytosolic DNA receptors were required for the response to R:D60, FL-DC cultures from mice lacking both Trif and MyD88 (*MyD88*$^{-/-}$/*Trif*$^{-/-}$) were transfected with R:D60. As shown in Figure 6.23 **a**, *MyD88*$^{-/-}$/*Trif*$^{-/-}$ FL-DCs failed to produce cytokines in response to R:D60. The cytokine response to ssDNA60 was also abolished.

Phenotypic activation in response to R:D60 was impaired, as shown for CD40 and CD86 (Figure 6.23 **b**). However, there appeared to be a difference between the two DC subtypes; in pDCs, expression was comparable to background levels (Lipofectamine alone); in cDCs, expression was reduced but not abolished. This may suggest that detection of R:D60 by cDCs could also occur in a MyD88-/Trif-independent manner, which fails to stimulate cytokine production. Further replicate experiments would be required to establish if this was significant.

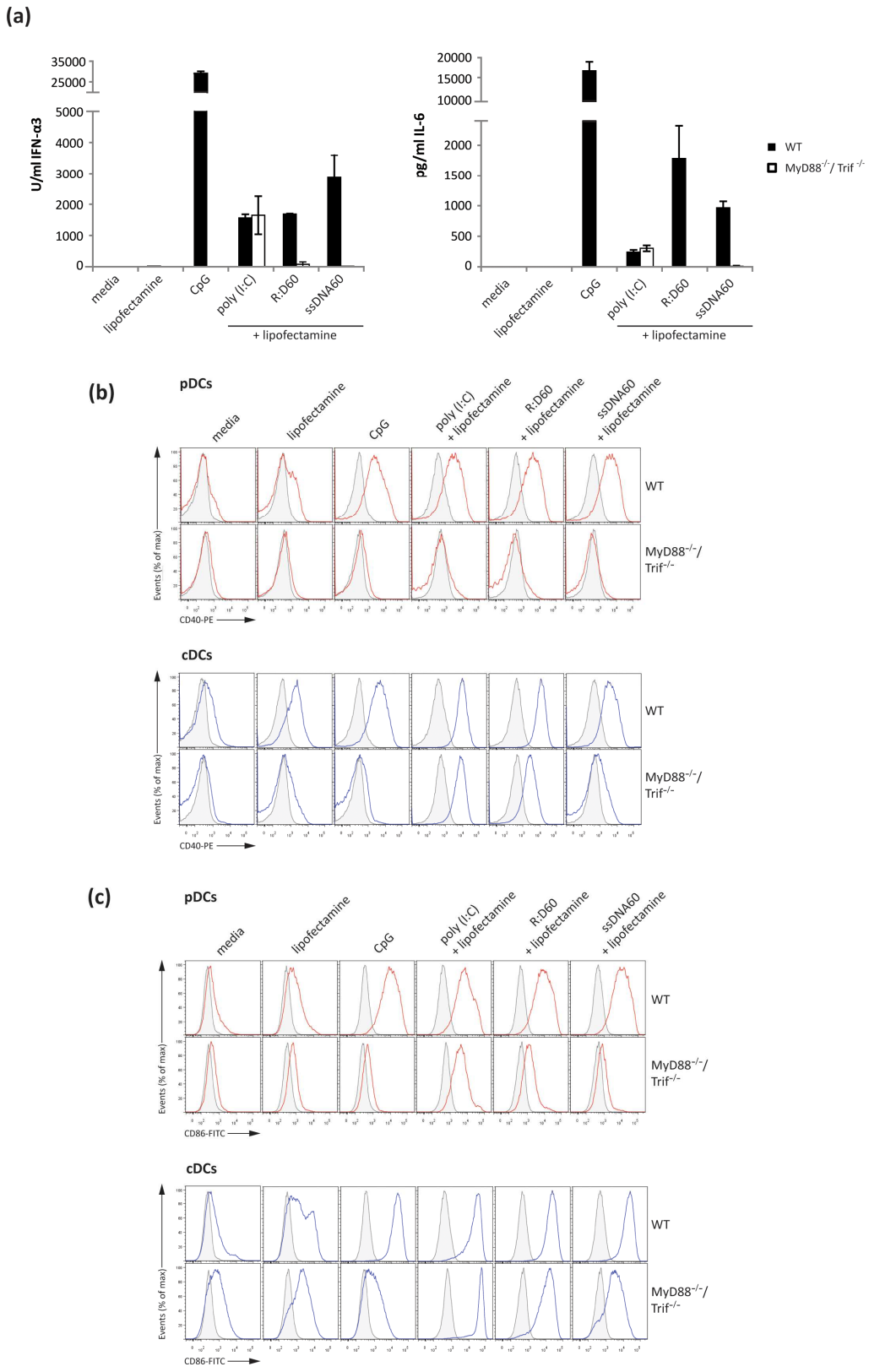


Figure 6.23. Trif and/or MyD88 is required for the response of FL-DCs to transfected R:D60.

(a) IFN-α and IL-6 production in response to R:D60 is abolished in *MyD88*$^{-/-}$/*Trif*$^{-/-}$. FL-DCs from *MyD88*$^{-/-}$/*Trif*$^{-/-}$ and age/sex matched WT controls were stimulated with 1 µg/ml CpG

ODN or were transfected with 1 μg/ml each of poly (I:C), R:D60 and ssDNA60 using Lipofectamine LTX. Supernatants and cells were analysed after 18 h. Data are from one experiment performed in duplicate ± s.d. **(b)** CD40 and **(c)** CD86 expression was reduced in *MyD88*$^{-/-}$/*Trif*$^{-/-}$ FL-DCs. Isotype controls are shown in grey, pDCs in red and cDCs in blue. Data are from one experiment, representative of duplicate wells

6.6.8.3. Reducing RNase H2 activity does not affect the cytokine response to R:D60

Finally, preliminary experiments to determine if RNase H2 plays a role in the response of FL-DCs to R:D60 were performed. The induction of type I IFN production by DNA via RIG-I-mediated signalling is indirect; the DNA is first recognised by RNA pol III in the cytosol and transcribed into immunostimulatory RNA (Ablasser *et al.*, 2009; Chiu *et al.*, 2009). Similarly, it is possible that R:D60 does not directly stimulate innate immunity in FL-DCs but instead requires initial detection and processing by RNase H2. Cleavage of R:D60, which was shown to be a substrate for RNase H enzymes in Chapter 5, leaves ssDNA and short polyribonucleotide products of digestion, which are immunostimulatory in FL-DCs (Figure 6.15). Consequently, FL-DCs with reduced levels of RNase H/H2 activity might be unable to mount a cytokine response to transfected R:D60.

An alternative hypothesis may be that RNase H2, in combination with other nucleases such as the ssDNA exonuclease Trex1, is required to degrade most of the R:D60, and so limit the extent of an immune response induced directly by the RNA:DNA hybrid. Expression of the *Rnaseh2* genes and *Trex1* mRNA can be detected in FL-DCs (Chapter 3, Figure 3.4). Preliminary analysis of type I IFN mRNA expression in FL-DCs transfected with R:D60 suggests that levels fall over time (from 6 h to 18 h) (Figure 6.12 and Figure 6.20), which could be due to processing of the hybrid. If this hypothesis was true, cells with reduced RNase H/H2 activity would produce increased levels of type I IFNs after 18 h.

To investigate this further, FL-DCs cultures from a mouse carrying the hypomorphic mutation Rnaseh2bA177T-neo/Rnaseh2A177T-stop (n/s) and a Rnaseh2bA177t-neo heterozygous littermate (n/+) were transfected with R:D60 and cytokine production measured by ELISA. Although RNase H2 activity levels in FL-DCs have not yet been quantified, activity is reduced by approximately 90% from wild type activity in MEFs established from n/s embryos. (see Chapter 3, Section 3.5.3). The relative proportions of pDCs and cDCs as determined by flow cytometry were the same for each mouse. FL-DCs generated from both mice produced approximately the same amount of IFN-α and IL-6 (Figure 6.24). This suggests that RNase H2 activity does not affect the response to R:D60, although further experiments are required to quantify this against wild type FL-DCs, as n/+ MEFs have approximately a 60% reduction in RNase H2 activity (see Chapter 3, Figure 3.11).

Finally, to establish if other immune cells from RNase H2-deficient mice can respond to RNA:DNA hybrids, bone marrow derived macrophages from the same mouse strains were were transfected with R:D60. However, as with wild type macrophages (see Section 6.4), R:D60 did not induce a cytokine response in these cells (Figure 6.25).

Together, this preliminary data suggested that RNase H2 is unlikely to be required as a sensor for RNA:DNA hybrids, as the cytokine response to R:D60 in FL-DCs from mice with a 90% reduction in RNase H2 activity is intact (Figure 6.24). Furthermore, in RNase H2-deficient cells, transfected R:D60 did not induce a cytokine production in a cells that are normally unresponsive to R:D60 (Figure 6.25).

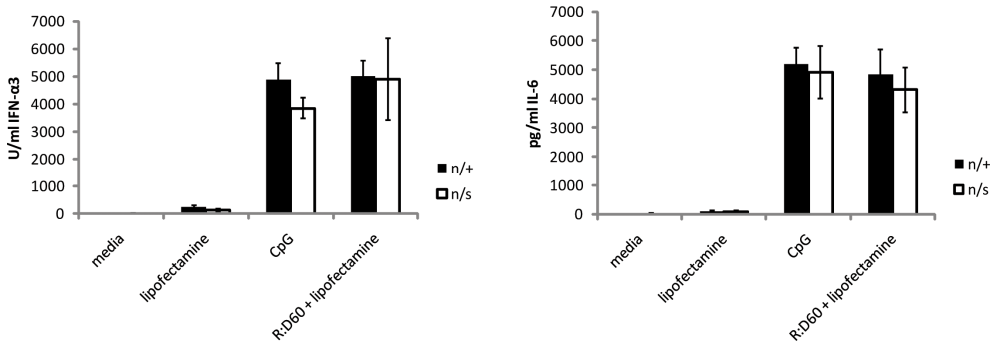


Figure 6.24. FL-DCs with reduced RNase H2 activity produce cytokines in response to transfected R:D60

FL-DC cultures from a Rnaseh2bA177T-neo/Rnaseh2bA177T-stop (n/s) mouse and a Rnaseh2A177Tneo heterozygous litter mate (n/+) were stimulated with 1 µg/ml CpG ODN or transfected with 1 µg/ml R:D60. After 18 h secreted cytokines were detected by ELISA. Data are from one experiment ± s.d. of triplicate ELISA wells.

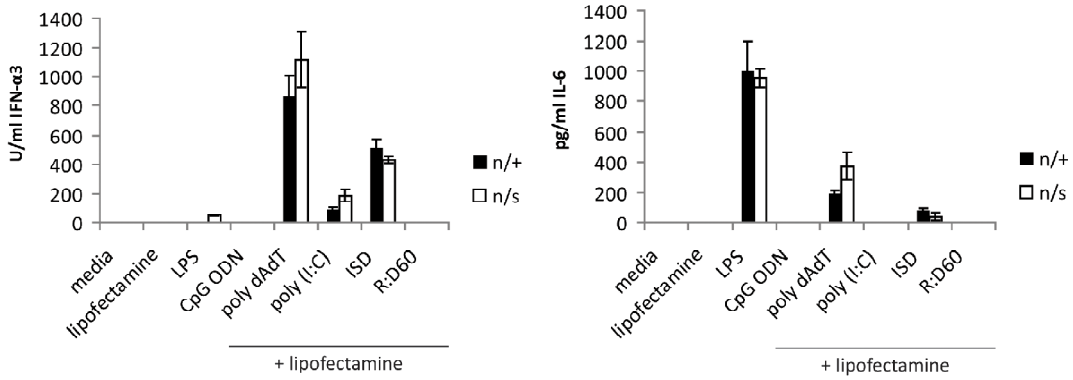


Figure 6.25. Macrophages deficient in RNase H2 activity do not produce cytokines in response to transfected R:D60.

Bone-marrow derived macrophage cultures were stimulated with PRR ligands for 18 h. Secreted cytokines were measured by ELISA. Rnaseh2bA177T-neo/Rnaseh2bA177T-stop compound heterozygote (n/s) and littermate Rnaseh2bA177T-neo heterozygotes (n/+) macrophages were stimulated with 100 ng/ml LPS or transfected with 1 µg/ml each of CpG ODN, poly (dA:dT), poly (I:C), 45 bp ISD and R:D60. Representative of two independent experiments, ± s.d. of triplicate ELISA wells.

6.7. Discussion

In this chapter, the immunostimulatory potential of RNA:DNA hybrids were tested *in vitro*. A 188 bp hybrid containing a 5′-ppp RNA strand induced type I IFN production in a number of cell types. In contrast, a 60 bp RNA:DNA hybrid without a 5′-ppp elicited a robust type I IFN response exclusively in pDC-containing FL-DC cultures.

6.7.1. Internalisation of RNA:DNA hybrids

Both hybrids required transfection into cells using the cationic-lipid transfection reagent Lipofectamine in order to stimulate a cytokine response (Figures 6.2 and 6.9). Studies using a fluorescently-labelled 18 bp hybrid revealed that this hybrid did not enter cells in the absence of transfection reagents (Figure 6.1). Therefore, it is likely that the absence of an immune response to exogenous R:D188 and R:D60 is a consequence of their failure to enter the cell. Although RNase H activity appears to be exclusively intracellular (Appendix 12), secreted nucleases, which include members of the RNase A superfamily (Sorrentino, 1998) and DNase I (Nadano *et al.*, 1993), may act to degrade free RNA:DNA hybrid before it enters the cell. The complexing of RNA:DNA hybrids to Lipofectamine via electrostatic interactions may protect from nuclease digestion. All of the hybrids used for the *in vitro* studies described in this chapter consist of fully PD-containing nucleic acids, therefore it would be informative to investigate the ability of nuclease resistant hybrids to enter a cell in the presence and absence of Lipofectamine, and ultimately to stimulate an immune response, for example, R:D20(PS) or R:D12 (Table 5.1).

Alternatively, it is possible that cells are unable to internalise extracellular RNA:DNA hybrids. The TLR9 ligand CpG ODN used as a positive control in the experiments throughout this chapter is partially PS-protected. It did not require transfection to stimulate a cytokine response in FL-DC cultures (Figure 6.9). Such CpG ODNs enter cells via endocytosis, mediated through non-specific interactions with cell surface proteins (Hacker *et al.*, 1998). This delivers them directly into the endosomal pathway, where they engage with TLR9 and stimulate cytokine

expression (Latz *et al.*, 2004). Cellular uptake of CpG ODNs is thought to be enhanced by the presence of PS-linkages (Sester *et al.*, 2000), which in addition to conferring increased nuclease resistance, lead to the formation of multimeric aggregates which are more readily endocytosed (Blasius and Beutler, 2010). However, many other nucleic acid ligands for PRRs of the innate immune system require complexing with transfection reagents or protein components to enter cells and engage with their receptors. Internalisation of un-complexed mammalian DNA is inefficient and consequently, innate immunity is not stimulated (Yasuda *et al.*, 2005b). However, intracellular targeting using transfection reagents, host proteins or autoantibodies localises self-DNA to cellular compartments where it can stimulate immunity via TLR9-dependent and independent pathways (Lande *et al.*, 2007; Leadbetter *et al.*, 2002; Yasuda *et al.*, 2005b). Similarly, intracellular targeting of both synthetic and self-RNA is necessary for stimulation of innate immunity both via TLR- and RLR-mediated pathways (Diebold *et al.*, 2004; Ganguly *et al.*, 2009; Heil *et al.*, 2004; Hornung *et al.*, 2006; Lau *et al.*, 2005; Pichlmair *et al.*, 2006; Vollmer *et al.*, 2005; Yasuda *et al.*, 2007). Therefore, it is likely that RNA:DNA hybrids require similar targeting to intracellular compartments to stimulate their receptors. The hybrids hypothesised to accumulate in RNase H2-deficient cells are endogenous hybrids, therefore would be translocated to the correct cellular location internally. This inability of RNA:DNA hybrids to stimulate cytokine production in the absence of transfection reagents reinforces the hypothesis that autoimmunity in AGS is likely to be driven through cell-intrinsic detection of nucleic acids (Stetson *et al.*, 2008).

6.7.2. The stimulation of type I IFN production by R:D188

Transfection of R:D188 into HeLa cells, L929 cells and primary MEFs stimulated a type I IFN response (Figures 6.2 and 6.3). In contrast, short hybrids failed to stimulate a response in the same cells. As described in Chapter 5, R:D188 differed from the other hybrids used in these experiments not only by length, but by sequence and the presence of a 5′-ppp on the RNA strand. In addition to the sensing of certain viral RNAs, RIG-I can detect transfected synthetic 5′-ppp dsRNA, in a largely sequence-independent manner (Schlee *et al.*, 2009b). Signalling by RIG-I is a two

step process, requiring detection of a 5′-ppp and RNA-dependent translocation of the protein (Myong *et al.*, 2009) (also discussed in Chapter 1). The structural similarities of RNA:DNA hybrid helices to those of dsRNA suggest that this mechanism could also occur for RNA:DNA hybrids. This hypothesis is strengthened by the evidence that RIG-I can translocate along short (< 50 bp) RNA:DNA duplexes (Myong *et al.*, 2009). In addition, R:D188 was generated from a viral sequence. Therefore, it could be envisaged that RIG-I is a candidate receptor for R:D188, although this does not exclude the stimulation of immunity via other sensors of dsRNA, including MDA5 and TLR3.

However, as described in Chapter 5, the purity of R:D188 is questionable. The use of reverse transcription to synthesise the DNA strand could potentially generate single-stranded contamination by separation of the template from the newly synthesised complementary ssDNA, or by incomplete transcription, thereby leaving ssRNA (Ma and Crouch, 1996). It was not possible to generate sufficient hybrid using the alternative annealing method described in Chapter 5 to perform *in vitro* testing, therefore reverse transcribed R:D188 was used. In an attempt to determine if the response to R:D188 was hybrid-specific, nuclease digested R:D188 was transfected into cells (Figure 6.2). However, RNase H digestion of the hybrid did not abrogate the type I IFN response. To determine if this could be attributed to the presence of contaminating nucleic acids in the preparation of R:D188, DNase I digestion of R:D188 was performed but even though it was used at low concentrations, DNase I did result in some cleavage of the hybrid (Figure 6.2 **c**). The induction of type I IFN production in both HeLa and L929 cells was increased when stimulated with DNase I-digested R:D188 (Figure 6.2 **d**). This implies that short ODN degradation products of DNase I, whether derived from contaminating DNA or the reverse strand of the hybrid, could be immunostimulatory in these cells. In addition, cleavage of the DNA strand of an RNA:DNA hybrid could release 188 nt 5′-ppp ssRNA, which itself is potentially immunostimulatory when transfected into cells. Therefore, the possibility exists that nuclease cleavage of RNA:DNA hybrids releases immunostimulatory degradation products.

Dot blot analysis of R:D188 using an antibody for ssNA suggested the presence of ssNA within this hybrid (Chapter 5, Figure 5.11). Digestion of R:D188 with RNase A in a buffer containing high salt levels should prevent cleavage of the hybrid whilst removing any ssRNA contaminants. RNase A treatment did not degrade the hybrid as determined by agarose gel electrophoresis (Figure 6.2 **d**). However, when transfected into cells, the amount of IFN-α produced by L929 cells decreased slightly, although this was not statistically significant (Student's *t*-test, $p = 0.572$). Digestion of R:D188 with RNase A combined with RNase H also failed to abolish the type I IFN response, indicating that 188 nt ssDNA or the products of RNaseA/H digestion are also immunostimulatory. Therefore, it was not possible to verify that the type I IFN response induced by transfected R:D188 is due to hybrid rather than the presence of contaminating ssDNA or degradation products. For this reason, R:D188 was not used in the subsequent experiments stimulating primary immune cells.

6.7.3. The immunostimulatory ability of short (≤ 60 bp) hybrids

The fractionation of a 60 bp $r(GU)_{30}$:$d(CA)_{30}$ duplex using FPLC successfully separated the double-stranded hybrid from ssNA (Chapter 5, Figure 5.4). Although potentially less physiologically relevant than a long hybrid produced by IVT and reverse transcription of a viral sequence, the greater confidence in the purity of R:D60 compared to R:D188 prompted further studies using this hybrid. Short hybrids did not stimulate a cytokine production in HeLa or L929 cells (Figure 6.2), however it is interesting to note that certain transformed cell lines are deficient in the use of certain PRR signalling pathways, specifically the cytosolic response to ISD (Stetson *et al.*, 2008; Stetson and Medzhitov, 2006). Therefore, R:D60 was transfected into primary cells which are known to produce type I IFNs in response to transfected nucleic acids, but again it failed to induce expression of type I IFN genes in primary MEFs (Figure 6.3). Although MEFs can utilise TLRs and RLRs following intracellular stimulation with nucleic acids, the failure to respond to transfected R:D60 could indicate that a specialist PRR or signalling pathway is required for

recognition of this hybrid, which is absent in this cell type. Therefore, the response of antigen presenting cells of the immune system to R:D60 was investigated.

Transfected R:D60 failed to stimulate a type I IFN response in bone marrow-derived macrophages (Figure 6.4). This was in contrast to the response to the PRR ligands CpG ODN, poly (I:C) and poly (dA:dT) which all stimulated induction of type I IFN expression. Transfected R:D60 also failed to stimulate pro-inflammatory cytokine production and phenotypic activation of macrophages, even when used at 25 μg/ml (Figure 6.5 and Figure 6.6).

6.7.4. R:D60 stimulates innate immune activation in FL-DCs

R:D60 also failed to elicit an immune response in GM-CSF cDCs (Figures 6.7 and 6.8), however when transfected into FL-DCs, it potently stimulated a cytokine response (Figure 6.10). As illustrated in Figure 6.9, FL-DCs are a heterogeneous population which can be segregated based on B220 staining into $B220^{+}$ pDCs and $B220^{-}$ cDCs. Flow cytometry analysis revealed that both subgroups were phenotypically activated in response to transfected R:D60 (Figure 6.11). Furthermore, transfected R:D60 stimulated a robust type I IFN response in FL-DC culture (Figure 6.10).

6.7.5. The response to R:D60 is unlikely to be due to contaminating nucleic acids

During this thesis, I aimed to generate RNA:DNA hybrids that were free from contamination with other nucleic acid species, with the rationale that ssNA contaminants would themselves be immunostimulatory. To test this in FL-DCs, cells were transfected with 500 ng/ml or 1 μg/ml of ssRNA60 and ssDNA60. Both stimulated IFN-α production, and especially the ssRNA (Figure 6.15). This confirmed that the purification of R:D60 from contaminating nucleic acids was required, to prevent the introduction of immunostimulatory ssNA into cells. However, it was also necessary to show that the process of purifying R:D60 away

from ssNA did not introduce any contaminants which might induce an immune response in FL-DCs. Sham annealings, where one of the strands is omitted from the hybridisation mixture, were fractionated using FPLC. As shown in Figure 6.17, the ssNA-containing peaks eluted from the column at different fractions compared to the equivalent elution of the hybrid (Figure 5.4). PAGE analysis of the fraction confirmed that neither ssRNA nor ssDNA eluted from the column at the same point as the hybrid (Figure 6.17). To verify that the chromatography process did not introduce other contaminants, the equivalent R:D60 containing fractions were transfected into FL-DCs. These fractions failed to induce either cytokine production, expression of *Ifna1* or phenotypic activation above background levels (Figure 6.18). Therefore, it can be concluded that FPLC fractionation of RNA:DNA hybrids did not introduce immunostimulatory contaminants into the hybrid preparation and that the purification of R:D60 using this method resulted in hybrid free from ssNA contaminants.

6.7.6. The nature of immunostimulatory RNA:DNA hybrids

Two other hybrids were tested for their ability to activate innate immunity in FL-DCs; the commercially available poly (rA:dT) and the 20 bp hybrid generated in Chapter 5 (R:D20). Whilst poly (rA:dT) failed to stimulate a response, a preliminary experiment using the 20 bp hybrid was at least as effective as R:D60 at inducing cytokine production (Figure 6.20). These experiments raised questions regarding the nature of immunostimulatory RNA:DNA hybrids, specifically, are only certain types of hybrid able to stimulate a type I IFN response in FL-DCs? Parameters such as the length and structural conformation may all be crucial in determining the ability of an RNA:DNA hybrid to induce a response. In particular, the specific sequence content of the hybrid could be a major factor, as both immunostimulatory hybrids used during this thesis had a GU-rich RNA strand. Further work varying these parameters is required to refine the precise characteristics of immunostimulatory hybrids.

6.7.7. pDCs as the IFN-α producing cells in FL-DCs transfected with R:D60?

FL-DCs are heterogeneous populations of DCs consisting of cDCs and pDCs (Figure 6.9). Each subpopulation has a distinct pattern of TLR expression (Table 6.1). pDCs, which express only TLR7 and TLR9, are specialised for the production of large amounts of IFN-α in response to the binding of nucleic acid ligands to these receptors (as discussed in Chapter 1, Section 1.5.5.1). They are therefore likely to be the source of the majority of IFN-α produced by FL-DC cultures in response to stimulation with CpG ODN (Figure 6.10), although the cDCs present in FL-DC cultures also express TLR9. IFN-α production by pDCs is also known to be stimulated via TLR7 by transfected 20 nt ssRNA (Heil *et al.*, 2004). Combined with the data presented in this chapter, pDCs appear to be a good candidate for the production of IFN-α in response to transfected R:D60. Furthermore, R:D60 failed to induce cytokine production or phenotypic activation in GM-CSF DC cultures, which do not contain pDCs (Figures 6.7 and 6.8). Analysis of activation markers in FL-DC cultures showed that both pDCs and cDCs had upregulated expression of CD40, CD86 and CD8 following transfection of R:D60 (Figure 6.11). Such phenotypic activation of cDCs could be the consequence of initial cytokine production by pDCs, which subsequently activates cDCs via a paracrine effect.

Intracellular cytokine staining could be used to identify the IFN-α-producing cell type in the FL-DC population. To investigate the hybrid activation of these cells in a more refined way, FL-DC cultures could be sorted using FACS or magnetic beads to allow the stimulation of pDCs and cDCs separately. Furthermore, analysis of *ex vivo* pDCs and cDCs could be used to investigate if the response to R:D60 is exclusive to DCs generated by Flt3-L differentiation of bone marrow progenitors.

6.7.8. A Trif and/or MyD88-dependent cytokine response to transfected R:D60

As discussed, the requirement of intracellular targeting of R:D60 inside the cell to stimulate an immune response suggests that FL-DCs are unable to endocytose R:D60, in contrast to CpG ODN. A previous study investigating the mechanism by which nucleic acid-cationic lipid complexes are internalised by a human epidermoid carcinoma cell line revealed that uptake is dependent upon the endocytic pathway, following fusion of the complex with the cell membrane (Almofti *et al.*, 2003). This is in agreement with an earlier observation that cellular trafficking of DNA that has been introduced into human fibroblasts via either endocytosis, transfection or electroporation occurs via endosomes (Coonrod *et al.*, 1997). Therefore, it is conceivable that upon uptake into the cell, R:D60 passes through endosomes and lysosomes, where it could engage with one of the intracellular TLRs. However, transfection using Lipofectamine is also used to introduce RLR ligands into cells (Hornung *et al.*, 2006; Schlee *et al.*, 2009b) which are presumed to be delivered to a cytosolic location.

The use of FL-DCs generated from mice deficient in the adaptors for these nucleic acid receptors showed that IPS-1 was not required for cytokine production or phenotypic activation in response to transfected hybrid (Figures 6.21 and 6.22). In contrast, the cytokine response to transfected R:D60 was abolished in FL-DCs derived from mice lacking both the TLR3 adaptor Trif and the TLR7/TLR9/DHX9/DHX36 adaptor MyD88 (Figure 6.23). This further suggests that RLRs are unlikely to be involved in innate recognition of R:D60, however replicate experiments are required to confirm this.

Therefore, these experiments indicate that TLR-signalling is potentially required for the cytokine response to R:D60 in FL-DC cultures. Further experiments utilising cells from mice deficient in the intracellular TLRs will be required to establish precisely which receptor the response to R:D60 is dependent on. The unaltered response of IPS-$1^{-/-}$ FL-DCs further strengthens the case for the role of pDCs in the

detection of R:D60, as pDCs generally rely on the TLR system for the sensing on nucleic acids (Kato *et al.*, 2005). Therefore, it is possible that the receptor could be either TLR7 or TLR9, or possibly both in the form of a TLR7/9 heterodimer. These receptors are both widely expressed by other cell types; including cDCs and macrophages, therefore determining why the response is likely to be pDC-specific will also require further investigation. However, the recent characterisation of MyD88-dependent cytosolic sensors of DNA in human pDC cell lines (DHX9 and DHX36, Kim *et al.*, 2010) raises the possibility that non-TLR receptors may play a role in the detection of intracellular R:D60.

In summary, in this chapter I have shown that a 60 bp highly-purified RNA:DNA hybrid stimulates a type I IFN response when intracellularly targeted into FL-DCs by complexing to a cationic-lipid transfection reagent. The cytokine response was shown to be MyD88 and/or Trif-dependent. The specialised type I IFN-producing pDCs are strongly implicated as the cell responsible for the detection of R:D60.

Chapter 7. Final discussion

It has been hypothesised that deficiencies in nuclease activity lead to accumulation of nucleic acids which then activate an innate immune response. In the autoimmune disorder AGS, loss-of-function mutations in the RNase H2 complex could result in an accumulation of RNA:DNA hybrids. This thesis aimed to establish roles for the RNase H2 complex and its nucleic acid substrates in the development of autoimmunity.

7.1. Summary of the main findings of this thesis

In Chapter 3, expression profiling confirmed that the genes encoding the RNase H2 complex are widely expressed, and in particular are expressed at high levels in certain cells of the immune system. This could suggest that RNase H2 activity is required to perform specific functions in these cell types; however, this was not pursued further during this thesis.

In contrast to *TREX1* and *SAMHD1*, expression of the *RNASEH2* genes was not induced in response to type I IFN production in primary MEFs. This implies that the RNase H2 complex is unlikely to act as a regulator of the cytosolic innate immune response to self-nucleic acids, previously proposed as a cellular function of both TREX1 and SAMHD1 (Rice *et al.*, 2009; Stetson *et al.*, 2008). Consequently, it appears most likely that an accumulation of RNase H2 substrates is the cause of RNase H2-AGS. Therefore, I sought to investigate the ability of RNA:DNA hybrids to activate innate immunity *in vitro*.

The nature of endogenous RNase H2 substrates likely to accumulate in the absence of enzymatic activity has not been established. Therefore, the unique physiochemical properties of RNA:DNA hybrids were determined, relevant to the design and analysis of synthetic hybrid duplexes. In Chapter 5, I demonstrated that whilst it was possible to generate large quantities of synthetic RNA:DNA hybrids, the production

of highly purified hybrids was more challenging, as preparations frequently contained contaminating nucleic acids that may also have immunostimulatory potential. Therefore, I established techniques to generate sufficient quantities of a 60 bp RNA:DNA hybrid (R:D60) and verified that contaminating nucleic acids had been removed.

Finally, in Chapter 6 I investigated the immunostimulatory potential of RNA:DNA hybrids, including the highly purified R:D60, via the stimulation of cell lines and primary cells. I showed that R:D60 induced innate immune activation when transfected into FL-DC cultures, and therefore propose pDCs as a cell type able to detect particular intracellular RNA:DNA hybrids. The cytokine response to R:D60 was MyD88/Trif-dependent, indicating that RNA:DNA hybrids are likely to be detected by intracellular TLRs or by recently characterised MyD88-dependent cytosolic DNA sensors (Kim *et al.*, 2010). In summary, I have shown that an RNA:DNA hybrid can induce innate immune signalling, resulting the production of IFN-α.

7.2. RNA:DNA hybrids: a novel PRR ligand?

7.2.1. Detection of R:D60 by pDCs?

The immune response to the 60 bp RNA:DNA hybrid generated during this thesis exclusively occurred in pDC-containing cultures. As discussed in Chapter 6, pDCs are therefore likely to be the cell type responsible for the production of IFN-α in response to transfected R:D60, via a MyD88- or Trif-dependent pathway. For this reason, it could be conceived that detection of R:D60 is via TLR7 and/or TLR9 in intracellular compartments, as pDCs express high levels of both receptors and the downstream transcription factor IRF7 (Coccia *et al.*, 2004; Izaguirre *et al.*, 2003; Kerkmann *et al.*, 2003). Notably, both TLR7 and TLR9 have been implicated in the detection of self-nucleic acids and development of disease pathogenesis in patients with another autoimmune disorder, SLE (Guiducci *et al.*, 2009; Komatsuda *et al.*, 2008).

It has previously been shown that the cellular response by TLR9 to nucleic acids is dependent upon the correct intracellular compartmentalisation of the ligands (Guiducci *et al.*, 2006). R:D60 failed to elicit a response without transfection, therefore hybrids appear to require intracellular trafficking via endocytic pathways to encounter their endosomal, or cytosolic, sensor. The treatment of R:D60 transfected cells with compound such as chloroquine could be used to determine if innate recognition of R:D60 occurs via membrane-bound receptors in acidified endosomal compartments. However, detection of ligands by TLR9 in the non-acidified early endosomes has been reported, notably in pDCs (Guiducci *et al.*, 2006), and so this would not be excluded by chloroquine treatment.

The recent characterisation of two novel MyD88-dependent cytosolic DNA sensors in human pDC cell lines (Kim *et al.*, 2010) raises the possibility that R:D60 may be detected in a TLR-independent manner. Although nucleic acid sensing pathways utilising these receptors have not yet been characterised in murine pDCs, it would be of interest to determine the requirement for these helicases in the cytokine response to R:D60. This may be achieved by using siRNA to knockdown expression of each of these proteins in FL-DCs.

7.2.2. Are intact RNA:DNA hybrids detected by PRRs?

The immune response to R:D60 may not be dependent on the direct detection of an RNA:DNA hybrid by PRRs. It is possible its single-stranded components, ssRNA60 and ssDNA60, are detected instead. For instance, R:D60 could be unwound by cellular helicases upon entry to the cell, such as Pif1 (Boule and Zakian, 2007), or replicative helicases (Shin and Kelman, 2006), although such proteins are predominantly localised in the nucleus. It is instead more likely that the hybrid may be degraded by intracellular nuclease activity. The subsequent release of either single strand would stimulate innate immune signalling in FL-DCs (Figure 6.15). Transfected RNase H-treated R:D60, which contains ssDNA60, stimulates innate immunity (Figure 6.16). Furthermore, ssDNA60 and R:D60 both fail to stimulate a response in the absence of transfection (Figure 6.10 and Appendix 13). Therefore, it

is entirely conceivable that the response is due to ssNA rather than direct recognition of the hybrid.

To address this, the intracellular stability of the hybrid will need to be investigated. A fluorescent oligonucleotide annealed to a complementary oligonucleotide which is coupled to a quencher molecule (for example, the RNase H2 substrate for the RNase H activity assay) could be transfected into cells. If this substrate is unwound or degraded, the fluorescein molecule will be released from the quencher and a fluorescent signal generated, which could be detected using flow cytometry or fluorescence microscopy. Alternatively, a time course could be performed whereby total nucleic acid extracts are prepared from cells transfected with R:D60 to detect the presence of such nucleic acid intermediates, similar to the techniques used to define the mechanism of RIG-I mediated immune activation by poly (dA:dT) (Ablasser *et al.*, 2009; Chiu *et al.*, 2009). Detection of RNA:DNA hybrids using the S9.6 antibody could be used to establish if the hybrid is stable after entering the cells using the techniques discussed in Chapter 4, Section 4.2.

ssDNA60 induced IL-6 production when transfected at high concentrations into macrophages but transfected R:D60 did not (25 and 10 μg/ml, Figure 6.5). DNA complexed to Lipofectamine is known to stimulate pro-inflammatory cytokine production via both TLR9-dependent and -independent pathways in macrophages (Yasuda *et al.*, 2005a). The lack of cytokine response to R:D60 implies that the hybrid is unable to stimulate cytokine production via the same pathway as ssDNA60. Furthermore, preliminary data obtained using FL-DCs generated from mice with significantly reduced RNase H2 activity argues against the intracellular processing of RNA:DNA hybrids into ssDNA by nuclease activity (Figure 6.24). If this was the case, it would be expected that the the cytokine response to hybrids would be substantially diminished in these cells. Nevertheless, although it is important to define if RNA:DNA hybrids are themselves a PRR ligand, the finding that innate immune activation can be stimulated by hybrids, even indirectly, is highly relevant to AGS disease pathogenesis.

7.3. RNA:DNA hybrids and human disease

The data presented in this thesis suggest that unprocessed RNA:DNA hybrids are capable of triggering an innate immune response. Therefore, one *in vivo* requirement for RNase H2 activity is likely to be the suppression of endogenous RNA:DNA hybrid accumulation. Failure to do this may result in activation of innate immunity by self-nucleic acids and the development of autoimmunity. This supports the notion that RNA:DNA hybrids are a trigger of innate immune activation in AGS patients.

7.3.1. Cell intrinsic mechanism for immune activation

The requirement for transfection of R:D60 to stimulate immunity suggests that innate recognition of RNA:DNA hybrids is cell intrinsic. This is consistent with the hypothesis that hybrids likely to accumulate in RNase H2-deficient cells are of endogenous origin. Such hybrids could potentially be generated during a number of cellular processes including DNA replication, endogenous retroelement replication or transcription (Figure 7.1 **a**). Two independent studies have reported the accumulation of endogenous DNA in the cytosol of Trex1-deficient cells, although different conclusions were drawn regarding the source of this DNA (Stetson *et al.*, 2008; Yang *et al.*, 2007). Therefore, the identification of accumulating RNase H2 substrates in cells with reduced enzymatic activity would further strengthen this argument and define the source of intracellular nucleic acids in AGS.

The major site of IFN-α production in AGS is intrathecal (Lebon *et al.*, 1988). Therefore, if RNase H2 and TREX1 deficiencies lead to the activation of cell intrinsic autoimmunity, it can be hypothesised that their substrates must accumulate in cells of the brain. Immunostaining of post-mortem AGS patient brains suggested that astrocytes were likely to be the predominant cell type responsible for cytokine production in the CNS (van Heteren *et al.*, 2008). Consequently, to fully elucidate disease pathogenesis, it would be of particular interest to investigate if R:D60 could stimulate a cytokine response in primary astrocytes.

7.3.2. Cell extrinsic mechanism for immune activation

In AGS, IFN-α production can also be systemic; the cytokine is often present at elevated levels in the sera, in addition to the CSF (Lebon *et al.*, 2002). However, the cell type responsible for IFN-α production outside of the CNS is unknown.

In addition to the hypothesised cell intrinsic detection of hybrids described in Section 7.3.1, it is also conceivable that detection of cell extrinsic self-RNA:DNA hybrids could occur in AGS. The autoimmune disorder SLE, which shares substantial phenotypic overlap with AGS, is characterised by the occurrence of large numbers of autoantibodies, particularly against self-nucleic acids. The formation of immune complexes (ICs) with self-DNA and self-RNA are known to stimulate IFN-α production via TLR7 and TLR9 in pDCs from SLE patients (Vollmer *et al.*, 2005; Yasuda *et al.*, 2007). Furthermore, the occurrence of autoantibodies against poly (rA:dT) has been reported in a small proportion of SLE patients (16/118) (Talal and Gallo, 1972). Preliminary testing of SLE patient plasma using methods described in Chapter 4 has revealed the presence of autoantibodies capable of detecting R:D18 in the plasma of one patient (Appendix 14). However, there was also a strong presence of autoantibodies detecting dsRNA; therefore it is unclear if the autoantibodies are simply detecting the A-form-like helical conformation shared by hybrids and dsRNA.

If circulating anti-RNA:DNA hybrid autoantibodies in AGS patients are highly prevelant, this could suggest an alternative mechanism whereby RNA:DNA hybrids are released from cells, for example during apoptosis or necrosis, and form ICs which can activate IFN-α production by pDCs (Figure 7.1 **b**). This could provide an explanation for the presence of significantly elevated levels of IFN-α in the plasma of AGS patients, and may be responsible for the development of systemic features of AGS, for example, the chilblain-like lesions present on the skin. SLE patients have a decreased number of circulating pDCs, as a consequence of recruitment to affected tissues; however pDC counts in the blood of AGS patients have been demonstrated to be normal (Van Heteren *et al.*, 2007) which indicates that the mechanism of disease differs between the two disorders. Therefore, to further examine a role for

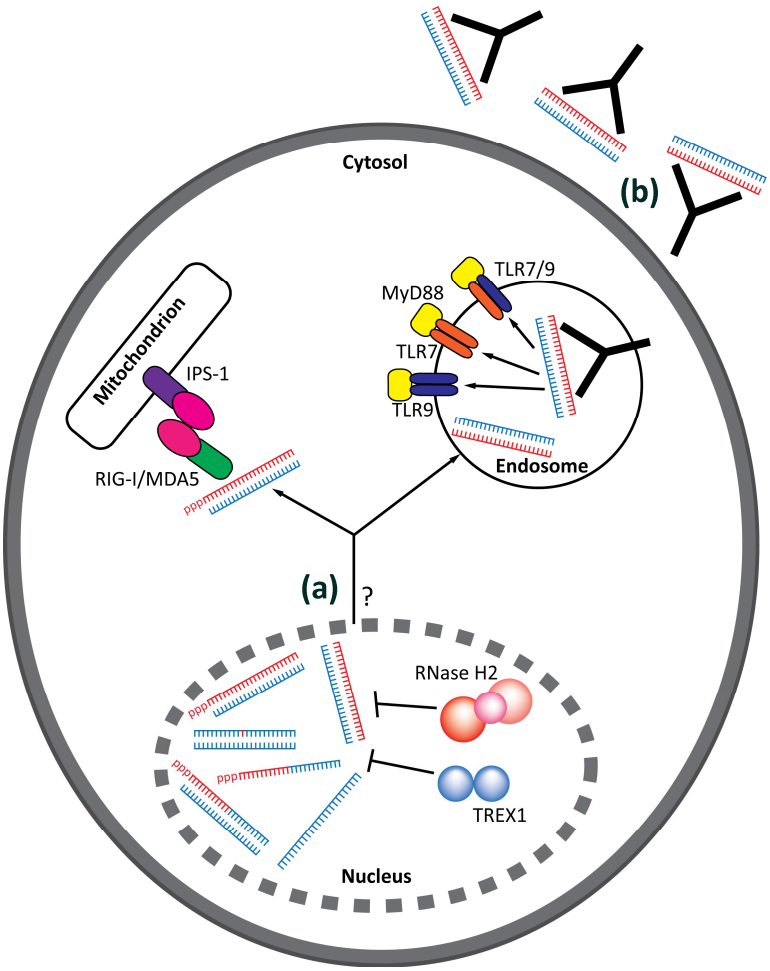


Figure 7.1. Intracellular detection of RNA:DNA hybrids.
Schematic representation of how RNA:DNA hybrids may be recognised by intracellular PRRs. **(a)** Cell intrinsic detection of hybrids. Nuclear RNase H2 and TREX1 may act to suppress accumulation of their nucleic acid substrates. RNA:DNA hybrid containing species could arise a consequence of R loop formation during transcription, activation of endogenous retroelements or during the generation of Okazaki fragments, amongst other processes. In the absence of RNase H2 and TREX1 activity, increased levels of these substrates are hypothesised to accumulate. Upon leaving the nucleus by an undefined mechanism, RNA:DNA hybrid-containing species could be detected in the cytosol, for example, by the RLRs RIG-I or MDA5 which utilise the adaptor protein IPS-1 to stimulate type I IFN production. It could be hypothesised that RNA:DNA hybrids with a 5'-ppp are potential ligands for RIG-I. In addition, it is possible that MyD88-dependent cytosolic DNA sensors (members of the DEAH/RHA helicase family) may play a role in the induction of type I IFNs as a consequence of cytosolic RNA:DNA hybrids. Alternatively, RNA:DNA hybrids may transported into the endosomal pathway, for instance by autophagy, and encounter receptive membrane-bound sensors. These sensors could include TLR7, TLR9 or a TLR7/9 heterodimer, which upon ligand binding stimulate downstream signalling and cytokine production. This does not exclude TLR3, however, pDCs do not express TLR3 therefore if they are the hybrid-sensing cells it is unlikely that TLR3 is required. **(b)** Cell extrinsic detection of hybrids. Circulating autoantibodies could form ICs with RNA:DNA hybrids released from cells, for example following apoptosis, and engage with Fc receptors on the surface of immune cells leading to antigen presentation to TLR7 and/or 9. This would

activate innate immune signalling via MyD88, and the subsequent production of type I IFN. Again, this does not exclude that TLR3 is involved; furthermore this mechanism could be independent of TLR-signalling. It is also conceivable that exogenous RNA:DNA hybrids could be internalised following complexing with peptides, for example LL37 or HMGB1.

pDCs in AGS, intracellular cytokine staining for IFN-α in pDCs purified from the blood of AGS patients could help to establish if they are the cell type responsible for the systemic IFN-α production in this disease. A murine model of RNase H2 AGS would be particularly useful in establishing if stimulation of innate immunity by RNA:DNA hybrids is via a cell intrinsic or extrinsic mechanism.

Studies using the Trex1-null mouse have shown that the induction of IFN production alone is not responsible for the development of a full inflammatory response, as the genetic ablation of RAG2 rescued the lethal autoinflammatory phenotype of these mice (Stetson *et al.*, 2008). This indicated a role for adaptive immunity in the development of disease. Therefore, obtaining a complete understanding of the pathogenesis of AGS will require investigation into how the initial innate immune response consequently activates and directs an adaptive response.

7.5. Summary

In summary, during this thesis I showed that a substrate of the RNase H2 enzyme complex was able to activate innate immune signalling in FL-DCs *in vitro*, in a MyD88 and/or Trif-dependent manner. This identifies RNA:DNA hybrids as a novel nucleic acid species able to stimulate IFN-α production, in contrast to a previous report which dismissed the capacity of hybrids to stimulate IFN production (Colby *et al.*, 1971). Activation of innate immunity by RNA:DNA hybrids appeared to be dependent on the sequence, length, and/or conformation of the hybrid and there appeared to be a cell type specificity in the response.

7.6. Further work

Further experiments are required to fully elucidate the immune signalling pathways activated in FL-DCs by the 60 bp RNA:DNA hybrid. Once the subtype of cell in which innate recognition of the hybrid occurs is identified, appropriate experiments utilising cells derived from mice deficient in PRRs can be performed, with TLR7 and/or TLR9 strong candidates for the receptors involved, although the involvement of the DEAH/RHA helicases DHX9 and DHX36 should also be investigated. Given some interspecies differences in PRR utilisation (Heil *et al.*, 2004), it will be important to establish if human DCs, specifically pDCs, can also produce IFN-α in response to R:D60.

The refinement of the exact nature of immunostimulatory hybrids is also of importance. Firstly, the generation and purification of 60 bp hybrids with different sequences to R:D60 will aid in determining if the response if sequence-dependent. Optimisation of the use of FPLC to purify hybrids of different lengths will permit investigation into whether there are any size constraints in the stimulation of immunity by hybrids, as well as clarifying whether purified long hybrids generated by reverse transcription, such as R:D188, are able to stimulate PRR pathways.

The RNA:DNA hybrids designed during this thesis represent only one type of nucleic acid species which could result from RNase H2-deficiency. As depicted in Figure 1.12, single-stranded nucleic acid flaps derived from Okazaki fragments consisting of a chimeric RNA-DNA strand, which may or may not be annealed to complementary DNA strand. Such 5′-ppp containing Okazaki fragment-like hybrids could be generated using RNA-primed PCR (Shibata *et al.*, 1995). Additionally, DNA duplexes containing single embedded ribonucleotides, such as those misincorporated by error-prone DNA polymerases, may also occur, also though the separation of a response to this type of substrate compared to full dsDNA would be particularly challenging.

AGS provides an example of an autoimmune disease which is mediated by inappropriate activation of the innate immune system. Nucleic acids are likely to be central to the mechanism of disease pathogenesis. The accurate characterisation of such nucleic acid ligands is essential for understanding how they activate immunity: failure to do so has previously caused confusion over the exact nature of specific PRR ligands (for example, synthetic ligands for RIG-I, as discussed by Schlee and Hartmann, 2010). The specific innate recognition pathways involved are not always immediately obvious (Ablasser *et al.*, 2009; Chiu *et al.*, 2009). Through the hypothesis that nucleic acids accumulate in AGS cells, I have been able to identify a novel nucleic acid species capable of activating the innate immune system. This is relevant to the understanding of nucleic acid-mediated autoinflammatory diseases, as well as the cellular response to exogenous nucleic acids.

Appendix 1. Oligonucleotides

(a) qRT-PCR: human

NAME	SEQUENCE (5' → 3')
CD14_R	CGAGGACCTAAAGATAACCGGC
CD14_R	GTTGCAGCTGAGATCGAGCAC
GAPDH_F	AGCCACATCGCTCAGACAC
GAPDH_R	GCCCAATACGACCAAATCC
IFNB1_F	CTGGCTGGAATGAGACTATTGTT
IFNB1_R	CAGGACTGTCTTCAGATGGTTTATC
IRF7_F	AAGGGCTTCCCCTGACTG
IRF7_R	TCTACTGCCCACCCGTACA
RNASEH2A_F	GAGAAAGAGGCGGAAGATGTTA
RNASEH2A_R	TCTTCCTGAGTCCCTCCTGA
RNASEH2B_F	GGGCTAATGTTTGTAAAACTGGTTA
RNASEH2B_R	TGCTGTAGACACATATTGAACAAGTAA
RNASEH2C_F	GGGACTCGAAGTGTCGTTTC
RNASEH2C_R	CCATCACGTATCCCACGAG
SAMHD1_F	CCCAAAGTATTGCTAGACGTGA
SAMHD1_R	TGCATTCCATAATCCATGTTG
TREX1_F	TGCCTTCTGTGTGGATAGCA
TREX1_R	GTGTTCTGAGGGGCTGCTT

(b) qRT-PCR: mouse

NAME	SEQUENCE (5' → 3')
BACTIN_F	CTAAGGCCAACCGTGAAAAG
BACTIN_R	ACCAGAGGCATACAGGGACA
IFNA1_F	GGAACAAGAGAGCCTTGACA
IFNA1_R	GAGGGTTGTATTCCATGCAG
IFNB1_F	GAACATTCGGAAATGTCAGG
IFNB1_R	ACTGTCTGCTGGTGGAGTTC
IL1B_F	CCCAACTGGTACATCAGCAC
IL1B_R	TCTGCTCATTCACGAAAAGG
RNASEH2A_F	TTGGTCTTGGGCACATTG

RNASEH2A_R	TGCATGTCTTTAACCTCAGCAC
RNASEH2B_F	AAGTACAGCTCAGAGAAGACATTGAA
RNASEH2B_R	CTTTTAGTGCCACAGTTTG
RNASEH2C_F	GGACTTCGACCGCCTTATC
RNASEH2C_R	CCAGACCCCACAAGGTGA
SAMHD1_F	TCAGTGAGCGAGATATACTCTGTGT
SAMHD1_R	AGAAAATGGCCCATGACCTA
TNF_F	CCCACTCTGACCCCTTTACT
TNF_R	TTTGAGTCCTTGATGGTGGT
TREX1_F	CAGGAGCAGGCGACCTAA
TREX1_R	CTCCTGTGATGGGGCTTC

(c) Genotyping of transgenic mice with mutations in *Rnaseh2b*

NAME	SEQUENCE (5' → 3')
LOXP_FORW	GATAGCTGACAAAGATAACTC
EXON7_REV	CCTGGAAACCTGACCACC
NEO_REV	CTGTCCATCTGCACGAGACT

(d) Oligonucleotide primers to generate RNA:DNA hybrids

188 bp hybrid

NAME	SEQUENCE (5' → 3')
T7_CMVPROM_F	TAATACGACTCACTATAGGCAAGTACGCCCCCTATTGAC
CMVPROM_R	GAAATCCCCGTGAGTCAAAC
CMVPROM_R_ECORI	GCCGAATTCGAAATCCCCGTGAGTCAAAC
CMVPROM_R_RTOVERHANG	AATTCGAAATCCCCGTGAGTCAAAC

51 bp hybrid

NAME	SEQUENCE (5' → 3')
T7_CMVPROM51_F	TAATACGACTCACTATAGGGTACATCAATGGGCGTGGATA
CMVPROM_R_ECORI	GCCGAATTCGAAATCCCCGTGAGTCAAAC

(e) Oligonucleotide primers to generate long ssDNA

NAME	SEQUENCE (5' → 3')
CMVPROM_F	CAAGTACGCCCCCTATTGAC
BIOTINCMVPROM_F	[BTN]CAAGTACGCCCCCTATTGAC
PHOSCMVPROM_F	[PHOS]CAAGTACGCCCCCTATTGAC
CMVPROM362_R	GTAGCGCTAGCGGATCTGAC
CMVPROM188_R	GAAATCCCCGTGAGTCAAAC

Appendix 2. Flow cytometry antibodies

FACSAriaII

Antibody	Label	Clone	Source	Isotype	Host	Dilution
CD11b	APC	M1/70	eBioscience	IgG2b	Rat	1:8
CD11c	FITC	N418	eBioscience	IgG	Armenian hamster	1:20
CD40	APC	1C10	eBioscience	IgG2a	Rat	1:8
CD80	PE	16-10A1	eBioscience	IgG	Armenian hamster	1:30
CD86	APC	GL1	eBioscience	IgG2a	Rat	1:30
F4/80	FITC	BM8	eBioscience	IgG2a	Rat	1:10
MHC Class II	PE	M5/114.15.2	eBioscience	IgG2b	Rat	1:100

FACSCantoII

Antibody	Label	Clone	Source	Isotype	Host	Dilution
B220/CD45R	Alexa-450	RA3-6B2	eBioscience	IgG2a	Rat	1:100
CD8	PE-Cy7	53-6.7	eBioscience	IgG2a	Rat	1:200
CD11b	PE	M1/470	eBioscience	IgG2b	Rat	1:400
CD11c	eFluor-780	N418	eBioscience	IgG1	Armenian hamster	1:100
CD11c	APC	N418	eBioscience	IgG1	Armenian hamster	1:200
CD24	FITC	M1/69	eBioscience	IgG2a	Rat	1:200
CD40	PE	3/23	BD Pharmingen	IgG2a	Rat	1:200
CD80	APC	16-10A1	eBioscience	IgG2	Armenian hamster	1:200
CD80	Alexa-488	16-10A1	eBioscience	IgG2	Armenian hamster	1:200
CD86	Alexa-488	GL-1	Biolegend	IgG2a	Rat	1:200
CD86	Alexa-700	GL-1	Biolegend	IgG2a	Rat	1:200
MHC Class II	PerCP-Cy5.5	M5114	Biolegend	IgG2b	Rat	1:1200

Appendix 3. Oligonucleotides annealed to produce RNA:DNA hybrids

NAME	SEQUENCE (5' → 3')
SSRNA12	AGCCUGGGAGCU
SSDNA12	AGCTCCCAGGCT
SSRNA18	GAUCUGAGCCUGGGAGCU
SSDNA18	AGCTCCCAGGCTCAGATC
SSRNA20	GUGUGUGUGUGUGUGUGUGU
SSDNA20	ACACACACACACACACACAC
SSRNA20(PS)	GCCCGUCUGUUGUGUGACUC
SSDNA20(PS)	GAGTCACACAACAGACGGGC
SSRNA60	GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGU
SSDNA60	ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC

Appendix 4. Polyacrylamide gel recipes

1. Native PAGE

	8%	12%	15%	20%
40% acrylamide/bis (19:1) (ml)	10	15	18.75	25
20 X TBE (ml)	1.25	1.25	1.25	1.25
10% APS (ml)	0.5	0.5	0.5	0.5
TEMED (ml)	0.05	0.05	0.05	0.05
ddH_2O (ml)	38.2	33.2	29.45	23.2
Total (ml)	50	50	50	50
Migration of reference dyes				
XC (bp)	160	70	60	45
BPB (bp)	45	20	15	12

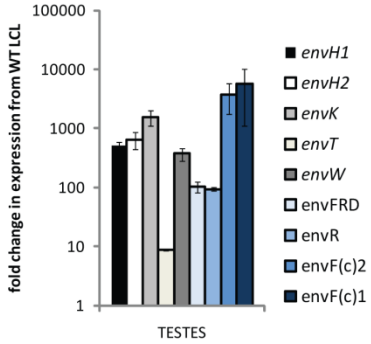
2. Denaturing PAGE

	6%	8%	12%	15%	20%
Urea (g)	25.2	25.2	25.2	25.2	25.2
40% acrylamide/bis (19:1) (ml)	9	12	18	22.5	30
20 X TBE (ml)	3	3	3	3	3
APS (ml)	0.4	0.4	0.4	0.4	0.4
TEMED (ml)	0.04	0.04	0.04	0.04	0.04
ddH_2O (ml)	26	23	17	12.5	5
Total (ml)	60	60	60	60	60
Migration of reference dyes					
XC (nt)	106	76		35	28
BPB (nt)	29	26		12	8

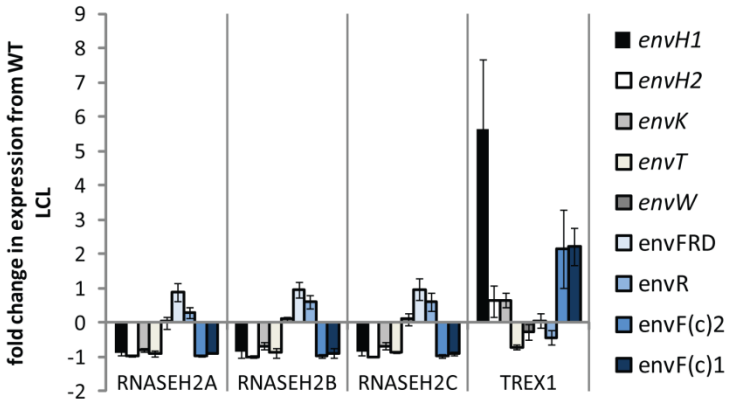
Appendix 5. HERV expression in AGS patient LCLs

RNA was extracted from four AGS patient LCLs and a normal control (WT LCL). cDNA was generated and qRT-PCR was performed using primers from de Parseval *et al*. 2003 to quantify the expression of HERV envelope genes. cDNA prepared from human testes RNA (from a FirstChoice® Human Total RNA Survey Panel, Ambion) was used to provide a positive control **(a)**, as *env* genes are expressed at high levels in the testes (de Parseval *et al.*, 2003). Normalised fold change in expression from WT LCL was calculated **(b)**. Data are from the mean fold change from one experiment ± s.d. of triplicate qPCR wells.

(a)



(b)



Appendix 6. Whole genome gene expression microarray analysis of homozygous Rnaseh2bA177T-neo MEFs versus wild type primary MEFs

Microarray data was validated by qRT-PCR, using the same RNA as the microarray samples and is represented as average fold change detected in three n/n versus three +/+ MEF lines. Expression levels of the four upregulated transcripts did not correlate with the array data. This suggested that the 13-fold increase in LOC383308 expression is likely to be a false positive, whilst the other genes show no change. The expression of a selection of transcripts downregulated in the mutant was verified by qRT-PCR (ND, not determined).

Up-regulated transcripts (total = 4) ($p \leq 0.05$, fold change ≥ 1.5)

Illumina ID	Gene symbol	Name	Fold increase (array)	*p*-value	Fold change (qRT-PCR)
ILMN_200537	LOC383308	Similar to ribosomal protein L29	12.98	2.09E-05	1.1 increase
ILMN_221019	RPL29	Ribosomal protein L29	2.11	7.27E-05	ND
ILMN_212510	4930455C21RIK	Hypothetical protein LOC76916	2.08	1.39E-02	0.0
ILMN_329373	WDR82	WD repeat domain 82	1.73	2.18E-04	1.1 decrease

Down-regulated transcripts (total = 48) ($p \leq 0.05$, fold change ≥ 1.5)

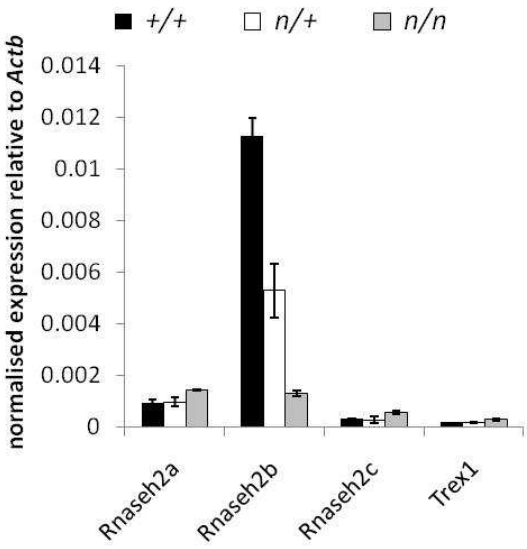
Illumina ID	Gene symbol	Name	Fold decrease (array)	*p*-value	Fold decrease (qRT-PCR)
ILMN_258184	ALOX5AP	arachidonate 5-lipoxygenase-activating protein	2.90	2.82E-02	3.8
ILMN_203628	E130112E08RIK	E130112E08RIK	2.57	2.92E-02	ND
ILMN_251042	CFP	complement factor properdin	2.33	3.45E-02	1.9
ILMN_210803	MRC1	mannose receptor, C type 1	2.31	3.76E-02	ND
ILMN_219967	C1QC	complement component 1, q subcomponent, C chain	2.30	4.27E-02	ND
ILMN_225389	CD68	CD68 molecule	2.22	2.18E-02	ND
ILMN_211748	C1QB	complement component 1, q subcomponent, beta polypeptide	2.20	3.46E-02	ND

ILMN_208807	MS4A6D	membrane-spanning 4-domains, subfamily A, member 6D	2.19	1.23E-02	ND
ILMN_239930	IFI27	interferon, alpha-inducible protein 27	2.19	4.66E-02	3.4
ILMN_213846	EMR1	EGF-like module containing, mucin-like, hormone receptor-like sequence 1	2.15	1.14E-02	ND
ILMN_223680	LMO2	LIM domain only 2	2.15	1.86E-02	ND
ILMN_220368	ARL11	ADP-ribosylation factor-like 11	2.11	7.21E-04	5.2
ILMN_212522	LAPTM5	lysosomal-associated protein transmembrane 5 [	2.03	3.31E-02	ND
ILMN_212075	MS4A7	membrane-spanning 4-domains, subfamily A, member 7	1.93	6.88E-03	ND
ILMN_217604	CCR5	chemokine (C-C motif) receptor 5	1.93	7.10E-03	3.4
ILMN_258659	CYTH4	cytohesin 4	1.90	4.01E-02	ND
ILMN_211039	CASP1	caspase 1	1.88	2.85E-03	2.0
ILMN_219201	CLEC4D	C-type lectin domain family 4, member d	1.88	4.53E-02	ND
ILMN_194028	P2RY6	pyrimidinergic receptor P2Y, G-protein coupled, 6	1.87	2.84E-02	ND
ILMN_192704	TREM2	triggering receptor expressed on myeloid cells 2	1.84	6.77E-03	ND
ILMN_2622209	2610207P08RIK	RNASEH2B	1.79	1.51E-03	8.5
ILMN_200650	MPEG1	macrophage expressed gene 1	1.76	3.35E-02	ND
ILMN_220410	NCKAP1L	NCK associated protein 1 like	1.76	3.49E-02	ND
ILMN_220005	LPXN	leupaxin	1.74	2.65E-02	ND
ILMN_189960	ENTPD4	ectonucleoside triphosphate diphosphohydrolase 4	1.71	2.22E-04	1.9
ILMN_212052	RNASEH2B	ribonuclease H2, subunit B	1.70	2.25E-03	ND
ILMN_186230	TYROBP	TYRO protein tyrosine kinase binding protein	1.70	8.46E-03	ND
ILMN_219342	CD93	CD93 molecule	1.68	1.70E-02	ND
ILMN_212868	FCGR4	Fc receptor, IgG, low affinity IV	1.68	1.99E-02	ND
ILMN_22116	MYO1F	myosin IF	1.66	2.03E-02	ND
ILMN_221982	NCF4	neutrophil cytosolic factor 4	1.65	4.35E-02	ND
ILMN_260026	DAB2	disabled homolog 2 (Drosophila)	1.65	4.48E-02	ND
ILMN_222602	BTK	Bruton agammaglobulinemia tyrosine kinase	1.64	3.73E-03	ND
ILMN_202115	IDB2	inhibitor of DNA binding 2	1.63	1.32E-02	ND

ILMN_194894	CTLA2B	cytotoxic T lymphocyte-associated protein 2 beta	1.63	2.14E-02	ND
ILMN_219885	CORO1A	coronin, actin binding protein 1A	1.61	2.74E-02	ND
ILMN_214875	CTSC	cathepsin C	1.61	3.80E-02	ND
ILMN_209055	LAT2	linker for activation of T cells family, member 2	1.60	1.33E-02	ND
ILMN_196998	LOC665032	similar to ribosomal protein L29	1.58	5.31E-04	2.1
ILMN_217284	DOK3	docking protein 3	1.57	3.94E-02	ND
ILMN_188421	PIK3CG	phosphoinositide-3-kinase, catalytic, gamma polypeptide	1.57	7.70E-03	ND
ILMN_193247	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	1.55	3.16E-02	ND
ILMN_193902	WAS	Wiskott-Aldrich syndrome homolog (human)	1.55	2.93E-02	ND
ILMN_218837	ARRB2	beta-arrestin 2	1.53	2.73E-02	ND
ILMN_220172	EVI2A	ecotropic viral integration site 2A	1.53	2.72E-02	ND
ILMN_192975	ZRANB3	zinc finger, RAN-binding domain containing 3	1.52	1.57E-02	ND
ILMN_187769	VAV1	vav 1 oncogene	1.50	1.63E-02	ND
ILMN_229771	TM6SF1	transmembrane 6 superfamily member 1	1.50	6.04E-03	ND

Appendix 7. Expression of AGS genes in homozygous Rnaseh2bA177T-neo primary MEFs

MEFs were established from wild type (+/+), *Rnaseh2bA177T-neo* heterozygous (n/+) and *Rnaseh2bA177T-neo*/*Rnaseh2bA177T-neo* homozygous littermate mate embryos. RNA was extracted from 3rd passage cells and qRT-PCR analysis of expression levels of the *Rnaseh2a*, *Rnaseh2b*, *Rnaseh2c* and *Trex1* genes determined. The primers for *Rnaseh2b* span exons 6 and 7. Expression data are the mean normalised relative expression for one embryo of each genotype ± s.d of triplicate PCR wells.



Appendix 8. Whole genome gene expression microarray analysis of homozygous Rnaseh2bA177T-neo/Rnaseh2b-A177T-stop primary MEFs versus wild type primary MEFs

Up-regulated transcripts (total = 173) ($p \leq 0.05$, fold change ≥ 2.0)

First 100 transcripts:

Illumina ID	Gene symbol	Name	Fold increase	*p*-value
ILMN_2950286	Ankrd1	ankyrin repeat domain 1 (cardiac muscle)	6.93	5.82E-03
ILMN_2937596	Ngf	nerve growth factor	5.12	8.91E-04
ILMN_2660233	Ngf	nerve growth factor	4.16	2.13E-03
ILMN_2531737	Dusp5	dual specificity phosphatase 5	3.70	1.13E-04
ILMN_2849449	Itga5	integrin alpha 5 (fibronectin receptor alpha)	3.45	5.69E-03
ILMN_2754985	Phlda1	pleckstrin homology-like domain, family A, member 1	3.36	1.09E-02
ILMN_2711461	Tgfb1	transforming growth factor, beta 1	3.32	3.00E-06
ILMN_1251114	Gltpd1	glycolipid transfer protein domain containing 1	3.32	3.90E-03
ILMN_2998658	Erdr1	erythroid differentiation regulator 1	3.21	4.88E-02
ILMN_1253304	Stmn2	stathmin-like 2	3.19	1.55E-04
ILMN_2759669	Scaf1	SR-related CTD-associated factor 1	3.19	1.82E-06
ILMN_2670172	Itga5	integrin alpha 5 (fibronectin receptor alpha)	3.15	9.75E-04
ILMN_2851288	Ngfr	nerve growth factor receptor (TNFR superfamily, member 16)	3.08	1.07E-04
ILMN_1216042	Apoe	apolipoprotein E	3.07	2.71E-02
ILMN_2692615	Tgm2	transglutaminase 2, C polypeptide	3.07	4.27E-03
ILMN_2661340	Slco4a1	solute carrier organic anion transporter family, member 4a1	3.07	8.27E-03
ILMN_2756704	9130213B05Rik	9130213B05Rik	3.00	8.00E-03
ILMN_2620406	Nptx2	neuronal pentraxin 2	2.92	8.70E-05
ILMN_1252295	Klra20	killer cell lectin-like receptor subfamily A, member 20	2.88	2.14E-03
ILMN_1225261	Uchl1	ubiquitin carboxy-terminal hydrolase L1	2.88	3.29E-04
ILMN_1230443	Tpm1	tropomyosin 1, alpha	2.87	1.13E-04
ILMN_2868220	Inhba	inhibin beta-A	2.86	1.82E-02
ILMN_2652187	Cux1	cut-like homeobox 1	2.86	2.30E-04
ILMN_2857957	Mgll	monoglyceride lipase	2.81	8.28E-04
ILMN_1236889	Selp	selectin, platelet	2.81	4.53E-04
ILMN_2624451	4933407C03Rik	4933407C03Rik	2.79	1.17E-06
ILMN_2760619	Tinagl	tubulointerstitial nephritis antigen-like 1	2.78	4.05E-05
ILMN_2715802	Hmga1	high mobility group AT-hook 1 [	2.74	2.33E-03
ILMN_2600744	Rgs16	regulator of G-protein signaling 16	2.73	4.25E-02
ILMN_2440530	Vat1	vesicle amine transport protein 1 homolog (T californica)	2.73	8.38E-05

ILMN_1228232	Cpe	carboxypeptidase E; similar to carboxypeptidase E	2.69	2.06E-03
ILMN_2522127	Myo1e	myosin IE	2.68	5.50E-07
ILMN_2901944	Col7a1	collagen, type VII, alpha 1	2.65	2.58E-02
ILMN_3153010	Slitrk5	SLIT and NTRK-like family, member 5	2.61	2.36E-03
ILMN_2730293	Pde1b	phosphodiesterase 1B, Ca2+-calmodulin dependent	2.60	1.09E-02
ILMN_2822579	Col4a2	collagen, type IV, alpha 2	2.60	1.80E-03
ILMN_2940713	Garnl4	GTPase activating RANGAP domain-like 4	2.60	1.01E-03
ILMN_1237606	Foxp4	forkhead box P4	2.59	1.21E-04
ILMN_1224034	Pde1b	phosphodiesterase 1B, Ca2+-calmodulin dependent	2.57	1.53E-02
ILMN_1226200	Sprr2k	small proline-rich protein 2K	2.57	4.54E-02
ILMN_2698449	Hbegf	heparin-binding EGF-like growth factor	2.56	2.55E-02
ILMN_2626389	Nomo1	nodal modulator 1	2.55	4.84E-05
ILMN_2955973	Smtn	smoothelin	2.55	2.83E-05
ILMN_2628757	Ehd1	EH-domain containing 1	2.53	6.42E-07
ILMN_2736514	Galntl4	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 4	2.52	2.25E-06
ILMN_2604521	Cotl1	adaptor protein complex AP-2, alpha 1 subunit	2.52	1.05E-03
ILMN_2944601	Pisd-ps3	phosphatidylserine decarboxylase, pseudogene 1	2.50	3.59E-03
ILMN_3143246	Ap2a1	adaptor protein complex AP-2, alpha 1 subunit	2.48	4.89E-05
ILMN_2436890	Dhrsx	dehydrogenase/reductase (SDR family) X chromosome	2.47	1.21E-06
ILMN_2854943	Gprc5a	G protein-coupled receptor, family C, group 5, member A	2.47	5.40E-05
ILMN_2855397	Hgs	HGF-regulated tyrosine kinase substrate	2.46	3.23E-06
ILMN_1236368	Orai1	ORAI calcium release-activated calcium modulator 1	2.46	1.64E-05
ILMN_1233243	Ivl	involucrin	2.45	1.56E-02
ILMN_1255422	Ccrn4l	similar to carbon catabolite repression 4 protein homolog	2.44	4.59E-02
ILMN_2713285	Fhl1	four and a half LIM domains 1	2.43	1.92E-02
ILMN_2675232	Klk8	kallikrein related-peptidase 8	2.43	1.39E-02
ILMN_1213989	Pmepa1	prostate transmembrane protein, androgen induced 1; similar to Nedd4 WW binding protein 4	2.42	2.42E-04
ILMN_3117381	Fhl1	four and a half LIM domains 1	2.40	2.79E-02
ILMN_2543630	Pisd-ps3	phosphatidylserine decarboxylase, pseudogene 3	2.39	1.87E-04
ILMN_3031402	Klra4	predicted gene 14378; similar to transforming growth factor, beta receptor III (betaglycan, 300kDa); mitogen-activated protein kinase kinase 7	2.39	3.16E-03
ILMN_2424299	Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	2.38	5.92E-03
ILMN_2598877	Slc7a8	solute carrier family 7 (cationic amino acid transporter, y+ system), member 8	2.36	2.28E-02
ILMN_2695181	Smtn	smoothelin	2.35	1.63E-04
ILMN_2675261	Taldo1	transaldolase 1	2.34	2.33E-05
ILMN_2729718	Psme3	proteaseome (prosome, macropain) 28 subunit, 3	2.33	7.31E-05
ILMN_2521893	Hmga2-ps1	high mobility group AT-hook 2, pseudogene 1	2.33	3.78E-03
ILMN_2767615	Atp1b1	ATPase, Na+/K+ transporting, beta 1 polypeptide	2.32	6.07E-04
ILMN_2836710	Cd151	CD151 antigen	2.31	1.14E-03
ILMN_2630605	Fscn1	fascin homolog 1, actin bundling protein (Strongylocentrotus purpuratus)	2.31	4.51E-04
ILMN_3105417	Bdnf	brain derived neurotrophic factor	2.31	1.80E-03
ILMN_1223551	Rrp9	RRP9, small subunit (SSU) processome component, homolog (yeast)	2.31	1.07E-04

ILMN_2656677	Fkbp8	FK506 binding protein 8	2.30	4.67E-06
ILMN_3116560	St3gal5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	2.30	5.65E-06
ILMN_2795078	Unc93b1	unc-93 homolog B1	2.29	3.49E-02
ILMN_1228501	Tob2	transducer of ERBB2, 2	2.28	3.15E-06
ILMN_1255175	Unc45a	unc-45 homolog A (C. elegans)	2.27	8.55E-07
ILMN_2511249	Akap2	A kinase (PRKA) anchor protein 2; paralemmin 2	2.27	2.69E-04
ILMN_1252995	Hist1h2be	histone cluster 1, H2be	2.26	1.88E-04
ILMN_1243862	Il11	interleukin 11	2.25	3.20E-02
ILMN_1253503	1190028F09		2.25	5.35E-03
ILMN_3148543	Src	Rous sarcoma oncogene	2.24	1.02E-05
ILMN_1254562	Atp13a2	ATPase type 13A2	2.24	2.33E-04
ILMN_1234072	Pdlim1	PDZ and LIM domain 1 (elfin)	2.24	2.10E-02
ILMN_2933394	Ppp1r10	protein phosphatase 1, regulatory subunit 10	2.24	2.89E-06
ILMN_2976129	Tinagl	tubulointerstitial nephritis antigen-like 1	2.23	4.13E-04
ILMN_2536389	LOC383897	LOC383897	2.23	1.09E-04
ILMN_3028575	Syngr1	synaptogyrin 1	2.23	1.53E-02
ILMN_1248181	Zbtb7a	zinc finger and BTB domain containing 7a	2.22	1.01E-02
ILMN_3110425	Myo1c	similar to nuclear myosin I beta; myosin IC	2.21	2.83E-05
ILMN_2808355	Coq5	coenzyme Q5 homolog, methyltransferase	2.21	8.35E-05
ILMN_2769567	F2rl1	coagulation factor II (thrombin) receptor-like 1	2.20	2.11E-02
ILMN_2592810	Map3k3	mitogen-activated protein kinase kinase kinase 3	2.20	5.52E-06
ILMN_2955919	Mcam	melanoma cell adhesion molecule	2.19	4.28E-02
ILMN_3026557	Rbpms	RNA binding protein gene with multiple splicing	2.19	7.29E-05
ILMN_2908056	Tmbim1	transmembrane BAX inhibitor motif containing 1	2.19	5.87E-08
ILMN_2648661	Cap1	CAP, adenylate cyclase-associated protein 1 (yeast)	2.19	3.97E-03
ILMN_2700689	Ibrdc3	ring finger protein 19B	2.19	9.35E-04
ILMN_2895000	Slc35c2	solute carrier family 35, member C2	2.18	1.78E-06
ILMN_2742455	Gtf2f1	general transcription factor IIF, polypeptide 1	2.18	4.99E-05
ILMN_2418811	2200002D01Rik	2200002D01Rik	2.16	1.96E-04

Down-regulated transcripts (total = 794) ($p \leq 0.05$, fold change ≥ 2.0)

First 100 transcripts:

Illumina ID	Gene symbol	Name	Fold decrease	P-value
ILMN_1258629	Col3a1	collagen, type III, alpha 1	10.89	8.12E-08
ILMN_2750515	Fos	FBJ osteosarcoma oncogene	8.97	1.92E-03
ILMN_2596346	Dcn	decorin	8.16	1.68E-03
ILMN_2662802	Ptx3	pentraxin related gene	8.15	1.48E-04
ILMN_2755660	LOC100044779	similar to prothymosin alpha	7.91	1.02E-07
ILMN_2494251	Tuba1a	tubulin, alpha 1A	7.42	6.10E-09
ILMN_1221290	LOC636952	LOC636952	6.95	6.88E-08
ILMN_2634317	Slc1a3	solute carrier family 1 (glial high affinity glutamate transporter), member 3	6.69	2.05E-03
ILMN_1217235	Fam114a1	family with sequence similarity 114, member A1	6.56	1.09E-09

ILMN_1253681	LOC433546	LOC433546	6.48	1.09E-09
ILMN_2906728	H19	H19 fetal liver mRNA	6.47	7.27E-04
ILMN_2773259	Lum	lumican	6.37	5.65E-03
ILMN_2955671	Slc38a2	solute carrier family 38, member 2	6.35	1.27E-07
ILMN_2504686	mt-Atp6	mt-Atp6	6.27	1.26E-08
ILMN_2638114	Ptn	pleiotrophin	6.27	5.05E-03
ILMN_2662803	Ptx3	pentraxin related gene	6.18	5.89E-04
ILMN_1247853	Zfp36l1	zinc finger protein 36, C3H type-like 1	6.14	8.09E-07
ILMN_2595612	Idh2	isocitrate dehydrogenase 2 (NADP+), mitochondrial	6.04	2.11E-06
ILMN_2683958	Col3a1	collagen, type III, alpha 1	5.98	2.51E-06
ILMN_2700233	Ccng2	cyclin G2	5.90	9.27E-10
ILMN_2419660	mtDNA_ND4L	mtDNA_ND4L	5.89	9.45E-07
ILMN_1235932	Pdgfra	platelet derived growth factor receptor, alpha polypeptide	5.88	2.88E-05
ILMN_3001540	Lum	lumican	5.81	2.23E-02
ILMN_1219335	Igfbp3	insulin-like growth factor binding protein 3	5.76	6.31E-03
ILMN_1232951	EG434404	ribosomal protein S6 pseudogene	5.68	5.74E-08
ILMN_2835423	Cfd	complement factor D (adipsin)	5.66	3.54E-04
ILMN_2497616	LOC100048480	LOC100048480	5.65	6.10E-09
ILMN_2643423	LOC100042773	LOC100042773	5.53	5.14E-08
ILMN_2668319	Hsp90aa1	heat shock protein 90, alpha (cytosolic), class A member 1	5.37	8.36E-08
ILMN_2768053	Supt16h	suppressor of Ty 16 homolog	5.30	1.63E-02
ILMN_1214602	Sfrp2	secreted frizzled-related protein 2	5.28	1.73E-02
ILMN_1249235	LOC668387	LOC668387	5.21	3.59E-08
ILMN_1216880	Emr1	EGF-like module containing, mucin-like, hormone receptor-like sequence 1	5.16	2.25E-05
ILMN_2426691	mt-Co2	mt-Co2	5.14	1.75E-08
ILMN_2835117	Ccl7	chemokine (C-C motif) ligand 7	5.08	1.13E-04
ILMN_1231689	Sfrp1	secreted frizzled-related protein 1	5.05	7.94E-04
ILMN_2846904	Mest	mesoderm specific transcript	5.02	8.68E-03
ILMN_2763245	Cxcl1	chemokine (C-X-C motif) ligand 1	5.01	2.02E-03
ILMN_2739760	Prelp	proline arginine-rich end leucine-rich repeat	5.00	3.22E-03
ILMN_2727503	Igfbp3	insulin-like growth factor binding protein 3	5.00	9.82E-03
ILMN_2530916	LOC225058	LOC225058	4.95	1.80E-08
ILMN_2832979	Gpc3	glypican 3	4.90	2.80E-03
ILMN_2669627	Vcan	versican	4.89	5.67E-08
ILMN_1230411	Kpnb1	karyopherin (importin) beta 1	4.77	8.31E-09
ILMN_2649502	Adam9	a disintegrin and metallopeptidase domain 9 (meltrin gamma)	4.75	9.27E-10
ILMN_2434853	mtDNA_ND2	mtDNA_ND2	4.74	1.84E-06
ILMN_2633229	Atp5a1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1	4.72	4.93E-09
ILMN_2485148	Vapa	vesicle-associated membrane protein, associated protein A	4.66	1.40E-09
ILMN_1249767	Thbd	thrombomodulin	4.63	1.74E-06
ILMN_2771815	Psmd14	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14	4.62	1.09E-09
ILMN_2677056	LOC100048622	LOC100048622	4.62	4.08E-09
ILMN_2903440	Hnrpdl	heterogeneous nuclear ribonucleoprotein D-like	4.59	5.46E-08
ILMN_2637639	Rps3	ribosomal protein S3	4.59	4.65E-07
ILMN_2662926	Egr1	early growth response 1	4.56	3.26E-05
ILMN_1230157	Rnd3	Rho family GTPase 3	4.53	2.88E-02
ILMN_1251119	Bmpr1a	bone morphogenetic protein receptor, type 1A	4.49	8.63E-09
ILMN_1256343	H19	H19 fetal liver mRNA	4.48	2.27E-03
ILMN_2527341	LOC381365	LOC381365	4.47	3.59E-08
ILMN_2513570	Rassf8	Ras association (RalGDS/AF-6) domain family (N-terminal) member 8	4.45	2.46E-09

ILMN_1258515	Paip2	polyadenylate-binding protein-interacting protein 2	4.44	1.09E-09
ILMN_2673936	Rpl18a	ribosomal protein L18A	4.41	1.16E-06
ILMN_1216231	Scoc	short coiled-coil protein	4.41	1.24E-04
ILMN_2607377	Itm2a	integral membrane protein 2A	4.40	8.63E-04
ILMN_2650008	Lgr5	leucine rich repeat containing G protein coupled receptor 5	4.38	4.49E-03
ILMN_1252832	Zc3h15	zinc finger CCCH-type containing 15	4.36	4.74E-08
ILMN_2674720	Ptges3	prostaglandin E synthase 3 (cytosolic)	4.31	2.58E-09
ILMN_2747959	Dcn	decorin	4.31	4.44E-02
ILMN_1243256	LOC675985	ribosomal protein S6 pseudogene	4.31	6.14E-09
ILMN_2725835	Ptp4a2	protein tyrosine phosphatase 4a2	4.24	7.65E-07
ILMN_2671601	Selk	selenoprotein K	4.23	4.31E-09
ILMN_2719973	Gpc3	glypican 3	4.23	1.18E-02
ILMN_2710166	Ddx3x	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3, X-linked	4.21	3.83E-09
ILMN_2702444	Csgalnact1	chondroitin sulfate N-acetylgalactosaminyltransferase 1	4.21	8.63E-09
ILMN_2757966	Pf4	platelet factor 4	4.20	5.61E-05
ILMN_2623216	Pgrmc1	progesterone receptor membrane component 1	4.16	1.95E-06
ILMN_2597332	1700123O20Rik	1700123O20Rik	4.16	3.29E-03
ILMN_1241536	LOC381649	LOC381649	4.13	4.08E-09
ILMN_2674966	Igsf10	immunoglobulin superfamily, member 10	4.12	3.88E-05
ILMN_2692723	Lpl	lipoprotein lipase	4.07	1.11E-02
ILMN_2604282	Sfrp1	secreted frizzled-related protein 1	4.06	1.56E-02
ILMN_3163159	LOC100041703	LOC100041703	4.06	1.97E-07
ILMN_2634338	1110002B05Rik	1110002B05Rik	4.02	2.32E-07
ILMN_2647628	Ppp1cb	protein phosphatase 1, catalytic subunit, beta isoform	3.94	1.01E-06
ILMN_2609998	Fxyd6	FXYD domain-containing ion transport regulator 6	3.94	1.46E-02
ILMN_1232081	Cald1	caldesmon 1	3.93	5.13E-07
ILMN_2495644	Vps29	vacuolar protein sorting 29	3.91	2.41E-08
ILMN_1253625	CNOT6	CCR4-NOT transcription complex, subunit 6	3.90	7.41E-08
ILMN_2621544	2700060E02Rik	2700060E02Rik	3.90	1.68E-04
ILMN_1259482	Pabpc1	poly(A) binding protein, cytoplasmic 1	3.88	2.20E-05
ILMN_1223428	Eif3s8	eukaryotic translation initiation factor 3, subunit C	3.87	1.02E-08
ILMN_1243212	Sparc	secreted acidic cysteine rich glycoprotein	3.87	4.18E-02
ILMN_1231873	Tmed2	transmembrane emp24 domain trafficking protein 2	3.87	4.31E-09
ILMN_1230620	Ywhaq	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide	3.86	1.60E-07
ILMN_1229547	Spon2	spondin 2, extracellular matrix protein	3.84	4.91E-03
ILMN_2680329	Isca1	iron-sulfur cluster assembly 1 homolog (S. cerevisiae)	3.84	5.93E-09
ILMN_1235619	Wasf2	WAS protein family, member 2	3.82	1.16E-07
ILMN_1230454	Hadhb	hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase trifunctional protein), beta subunit	3.81	8.43E-08
ILMN_2605465	Snx1	sorting nexin 1	3.76	7.96E-10
ILMN_2874853	Eef1b2	eukaryotic translation elongation factor 1 beta 2	3.76	1.87E-06

Appendix 9. Phenotypes of homozygous Rnaseh2bA177T-stop embryos used for microarray analysis

Embryos were dissected and phenotype analysis performed by Kirstie Lawson, MRC HGU, prior to RNA extraction. The RNA yield listed was the quantity recovered after DNase I treatment and clean-up using spin columns.

Litter number

E8.5	Number of WT embryos
1	1
2	2

E9.5	Number of WT embryos	Number of mutant embryos
1	0	1
2	1	1
3	2	0
4	0	1

Embryo	Stage	Observations	Total RNA yield
E8.5 (WT)			
1	TS13	9 somites. Embryo has begun to turn.	1.4μg
2	TS12	6/7 somites	0.8μg
3	TS12	6/7 somites	0.5μg
E9.5 (mutant)			
1	abnormal	Small embryo. Head is $^{1}/_{4}$ turned. Equivalent of 7 somite stage. Mesenchyme is opaque. Approx. 7/8 somites, all fused.	1.1μg
2	abnormal	Fairly normal E8.5+ anteriorly but embryo is opaque and hasn't turned. Truncated below 6th somite. Ventral allantois. Probably fused.	1.4μg
3	abnornal	Growth retarded and opaque. Brain and heart are equivalent of 7 somite stage. Head has turned. No visible somites. Rudimentary allantois. Amnion restricted and no large blood vessels in yolk sac.	1.4μg
E9.5 (WT)			
1	TS15	22 somites	3.8μg
2	TS16	22 somites	8.2μg
3	TS16	23 somites	8.7μg

Appendix 10. Whole genome gene expression microarray analysis of E9.5 homozygous Rnaseh2bA177T-stop embryos

E9.5 mutant versus E9.5 wild type

Up-regulated transcripts (total = 117) ($p \leq 0.05$, fold change ≥ 1.5)

Illumina ID	Gene symbol	Name	Fold increase	P-value
ILMN_2500276	Ccng1	Cyclin G1	6.54	1.37E-05
ILMN_2634083	Cdkn1a	Cyclin-dependent kinase inhibitor 1A (p21)	6.13	4.52E-07
ILMN_2710229	Ccng1	Cyclin G1	5.15	1.89E-04
ILMN_2846775	Cdkn1a	Cyclin-dependent kinase inhibitor 1A (p21)	4.14	4.52E-07
ILMN_2923607	Phlda3	Plekstrin homology-like domain, family A, member 3	4.12	1.45E-05
ILMN_2702233	Ccng1	Cyclin G1	4.12	3.35E-04
ILMN_2628567	Phlda3	Plekstrin homology-like domain, family A, member 3	3.97	5.98E-03
ILMN_2506012	Trp53inp1	Transformation related protein 53 inducible nuclear protein 1	3.46	2.30E-04
ILMN_2868280	Dcxr	Dicarbonyl/L-xylulose reductase	3.13	3.34E-02
ILMN_2846776	Cdkn1a	Cyclin-dependent kinase inhibitor 1A (p21)	3.05	1.27E-04
ILMN_2765454	Svop	SV2 related protein	2.95	6.15E-03
ILMN_2759108	Dcxr	Dicarbonyl/L-xylulose reductase	2.73	2.66E-02
ILMN_1250531	Slc19a2	Solute carrier family 19 (thiamine transporter), member 2	2.72	1.89E-04
ILMN_2764551	Cox6b2	Cytochrome c oxidase subunit VIb polypeptide 2	2.64	3.35E-04
ILMN_2418426	Zmat3	Wild-type p53-induced gene 1	2.54	2.17E-04
ILMN_2617256	Ms4a10	Membrane-spanning 4-domains, subfamily A, member 10 (	2.45	1.09E-02
ILMN_2717678	Muc13	Mucin 13, epithelial transmembrane	2.38	1.33E-02
ILMN_2695819	Ddit4l	DNA-damage-inducible transcript 4-like	2.37	1.89E-04
ILMN_1232336	1700007K13Rik	Hypothetical protein LOC69327	2.24	2.84E-02
ILMN_2879681	B4galnt2	Beta-1,4-N-acetyl-galactosaminyl transferase 2	2.24	6.15E-03
ILMN_1217442	Zfp365	Zinc finger protein 365	2.22	8.72E-03
ILMN_1228469	Apoc1	Apolipoprotein C-I	2.21	1.09E-02
ILMN_2518406	Vil1	Villin 1	2.20	3.65E-02
ILMN_2699466	Pla2g12b	Phospholipase A2, group XIIB	2.18	6.15E-03
ILMN_2615096	Dpp4	Dipeptidylpeptidase 4	2.15	4.11E-02
ILMN_3163053	Commd3	COMM domain containing 3	2.12	1.30E-04
ILMN_2599794	Apoc1	Apolipoprotein C-I	2.04	4.38E-02
ILMN_2483304	Lyve1	Lymphatic vessel endothelial hyaluronan receptor 1	2.04	2.59E-02
ILMN_2838317	Pqlc3	PQ loop repeat containing	1.98	1.41E-02
ILMN_2477136	2010003K15Rik	Hothetical protein LOC75606	1.96	1.85E-04
ILMN_2948945	Sesn2	Sestrin 2	1.94	1.96E-02
ILMN_1245909	4933402E13Rik	Hypothetical protein LOC74437	1.91	6.66E-03
ILMN_2497999	Vkorc1	Vitamin K epoxide reductase complex, subunit 1	1.91	3.61E-02
ILMN_1246289	Got1l1	Glutamic-oxaloacetic transaminase 1-like 1	1.90	5.98E-04
ILMN_2419185	Tuft1	Tuftelin 1	1.90	9.05E-03
ILMN_2647313	Klhl26	Kelch-like 26	1.89	8.52E-03
ILMN_2908065	Gtse1	G two S phase expressed protein 1	1.89	6.67E-03
ILMN_2850753	Mdm2	Transformed mouse 3T3 cell double minute 2	1.89	3.93E-02
ILMN_1219298	Art4	ADP-ribosyltransferase 4	1.89	6.15E-03
ILMN_2661650	Reep6	Polyposis locus protein 1-like 1	1.86	6.15E-03
ILMN_1238634	4933402E13Rik	Hypothetical protein LOC74437	1.85	8.44E-03
ILMN_2630428	Fmr1nb	Fragile X mental retardation 1 neighbor	1.85	1.33E-02
ILMN_3136561	Sparc	Secreted acidic cysteine rich glycoprotein	1.84	1.75E-02
ILMN_1223880	Tmprss2	Transmembrane protease, serine 2	1.84	8.72E-03

ILMN_2593134	Rnasek	Ribonuclease kappa	1.84	3.05E-02
ILMN_2728473	Serpina3m	Serine (or cysteine) proteinase inhibitor, clade A, member 3M	1.84	1.86E-02
ILMN_3162879	Stx3	Syntaxin 3	1.84	8.72E-03
ILMN_2920736	Bbc3	BCL2 binding component 3	1.82	2.58E-03
ILMN_2667805	Ak1	Adenylate kinase 1	1.81	1.02E-02
ILMN_2777427	Nostrin	Nitric oxide synthase trafficker	1.77	3.85E-02
ILMN_2698046	Stat3	Signal transducer and activator of transcription 3	1.77	6.15E-03
ILMN_3160626	Cideb	Cell death-inducing DNA fragmentation factor, alpha subunit-like effector B	1.77	2.76E-02
ILMN_1240445	Stard4	StAR-related lipid transfer (START) domain containing 4	1.77	3.05E-02
ILMN_2707541	Folr1	Folate receptor 1 (adult)	1.74	3.51E-02
ILMN_2641371	Lgals2	Lectin, galactose-binding, soluble 2	1.73	1.49E-03
ILMN_3002745	Vkorc1	Vitamin K epoxide reductase complex, subunit 1	1.73	3.93E-02
ILMN_1255479	Cldn2	Claudin 2	1.72	1.49E-03
ILMN_1235688	Cd151	CD151 antigen	1.72	2.96E-02
ILMN_2747857	Slc27a3	Solute carrier family 27 member 3	1.72	3.48E-03
ILMN_2980323	Cldn2	Claudin 2	1.71	3.05E-02
ILMN_2861123	Tmem150	Transmembrane protein 150A	1.71	2.71E-02
ILMN_2767825	2210011C24Rik	Hypothetical protein LOC70134	1.71	4.11E-02
ILMN_2908070	Gtse1	Two S phase expressed protein 1	1.71	1.49E-03
ILMN_2628594	Mttp	Microsomal triglyceride transfer protein	1.70	3.45E-02
ILMN_2759762	Stard10	START domain containing 10	1.70	4.29E-02
ILMN_1255438	Cpped1	Calcineurin-like phosphoesterase domain containing 1	1.70	4.38E-02
ILMN_1240630	Polk	DNA-directed DNA polymerase kappa	1.70	3.05E-02
ILMN_2492403	Cd151	CD151 antigen	1.70	1.86E-02
ILMN_2690575	Pcbd1	Pterin 4 alpha carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha	1.69	3.61E-02
ILMN_2595842	Def6	Differentially expressed in FDCP 6	1.68	6.58E-03
ILMN_2915833	BC025446	BC025446	1.68	8.52E-03
ILMN_2644936	Habp2	Hyaluronic acid binding protein 2	1.68	3.11E-02
ILMN_1239293	Sulf2	Sulfatase 2	1.66	4.60E-02
ILMN_2907655	Bax	Bcl2-associated X protein	1.66	8.72E-03
ILMN_2522693	Vkorc1	Vitamin K epoxide reductase complex, subunit 1	1.65	4.68E-02
ILMN_2476139	Tuba6	Tubulin alpha 6	1.65	2.36E-02
ILMN_2770429	Plod2	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	1.65	8.72E-03
ILMN_2774349	LOC100046401	Similar to SDR2	1.64	9.08E-03
ILMN_2915835	BC025446	BC025446	1.64	4.60E-02
ILMN_2604310	1810022C23Rik	Hypothetical protein LOC69123	1.64	1.27E-02
ILMN_1227282	Soat2	Acyl-CoA:cholesterol acyltransferase 2	1.64	2.01E-02
ILMN_2730651	Slc7a9	Solute carrier family 7, member 9	1.63	4.11E-02
ILMN_1226293	Slc39a4	Solute carrier family 39, member 4	1.63	4.38E-02
ILMN_1217102	Tpcn1	Two pore channel 1	1.62	1.52E-02
ILMN_2953677	Ei24	Etoposide induced 2.4	1.62	8.72E-03
ILMN_2655260	Ptp4a3	Protein tyrosine phosphatase 4a3	1.62	4.91E-02
ILMN_2775512	Ei24	Etoposide induced 2.4	1.62	4.38E-02
ILMN_1235811	Fam198b	Expressed in nerve and epithelium during development	1.61	6.68E-03
ILMN_2979190	Habp2	Hyaluronic acid binding protein 2	1.60	3.27E-03
ILMN_2992009	Lrdd	Leucine-rich and death domain containing	1.60	2.66E-02
ILMN_2621708	Abcc4	ATP-binding cassette, sub-family C (CFTR/MRP), member 4	1.60	1.47E-02
ILMN_2874554	Aldh4a1	Aldehyde dehydrogenase 4A1	1.59	4.12E-02
ILMN_1247811	Ass1	Argininosuccinate synthetase 1	1.59	1.09E-02
ILMN_1236522	Cbr1	Carbonyl reductase 1	1.58	2.87E-02
ILMN_2702102	D630023F18Rik	Hypothetical protein LOC98303	1.58	1.49E-03
ILMN_2806996	Creb3l3	cAMP responsive element binding protein 3-like 3	1.58	6.67E-03
ILMN_2671165	Krt23	Keratin 23	1.58	1.68E-02
ILMN_2665715	Gstt3	Glutathione S-transferase, theta 3	1.57	2.66E-02
ILMN_2715546	Gpx3	Glutathione peroxidase 3	1.57	2.40E-02
ILMN_2535430	LOC383576	LOC383576	1.56	3.37E-02
ILMN_3143577	Apaf1	Apoptotic protease activating factor 1	1.56	2.52E-02
ILMN_1248039	mtDNA_COXII	mtDNA_COXII	1.56	6.58E-03
ILMN_1249550	Fkbp9	FK506 binding protein 9	1.56	2.79E-02

ILMN_1232495	LOC100044177	Hypothetical protein LOC100044177	1.56	3.81E-02
ILMN_2775360	Mapkapk3	Mitogen-activated protein kinase-activated protein kinase 3	1.56	4.89E-02
ILMN_1220609	Ltb4r1	Leukotriene B4 receptor 1	1.55	4.67E-02
ILMN_1235474	Pdzk1	PDZ domain containing 1	1.55	2.84E-02
ILMN_1246764	Fam122b	Family with sequence similarity 122, member B	1.55	1.80E-02
ILMN_1247490	Arhgap23	Rho GTPase activating protein 23	1.54	2.66E-02
ILMN_2546272	Gng12	Guanine nucleotide binding protein (G protein), gamma 12,	1.54	2.33E-02
ILMN_1243433	Gm10080	Predicted gene 10080	1.54	2.84E-02
ILMN_2596611	Clic1	Chloride intracellular channel 1	1.54	2.15E-02
ILMN_2946760	Ankrd56	Ankyrin repeat domain 56	1.52	6.67E-03
ILMN_2631093	Cd81	CD81 antigen	1.52	6.15E-03
ILMN_2773485	Thnsl2	Threonine synthase-like 2	1.52	4.60E-02
ILMN_1219136	Vamp8	Vesicle-associated membrane protein 8	1.51	2.84E-02
ILMN_2606921	Klb	Klotho beta	1.51	4.60E-02

E9.5 mutant versus E9.5 wild type

Down-regulated transcripts (total = 29) ($p \leq 0.05$, fold change ≥ 1.5)

Illumina ID	Gene symbol	Name	Fold decrease	P value
ILMN_2622209	Rnaseh2b	Ribonuclease H2, subunit B	3.06	8.65E-07
ILMN_2622208	Rnaseh2b	Ribonuclease H2, subunit B	2.87	3.27E-07
ILMN_1233402	LOC100045981	Similar to synaptotagmin XI	2.40	2.48E-04
ILMN_1231030	Tcf15	Transcription factor 15	1.96	1.22E-05
ILMN_2738123	LOC100046518	Similar to Gprin1 protein	1.87	1.78E-04
ILMN_2828112	Igfbpl1	Insulin-like growth factor binding protein-like 1	1.84	8.29E-05
ILMN_1256680	Jakmip2	Janus kinase and microtubule interacting protein 2	1.78	2.86E-04
ILMN_1233535	6330407J23Rik	Hypothetical protein LOC67412	1.73	1.81E-06
ILMN_2976821	Fzd10	Frizzled 10	1.72	6.51E-06
ILMN_2584527	Mbd2	Methyl-CpG binding domain protein 2	1.65	1.03E-05
ILMN_1226157	Pik3r3	Phosphatidylinositol 3 kinase, regulatory subunit, polypeptide 3 (p55)	1.64	6.25E-05
ILMN_2565262	Lrig3	Leucine-rich repeats and immunoglobulin-like domains 3	1.62	8.34E-05
ILMN_2776554	Igfbpl1	Insulin-like growth factor binding protein-like 1	1.62	5.56E-05
ILMN_1247530	Llgl1	lethal giant larvae homolog 1	1.61	2.08E-04
ILMN_2657409	Rps18	Ribosomal protein S18	1.60	4.98E-05
ILMN_1220516	Wnt3a	wingless-related MMTV integration site 3A	1.59	3.74E-05
ILMN_1260073	D16Ertd472e	D16Ertd472e	1.59	1.01E-04
ILMN_2776909	Irx3	Iroquois related homeobox 3	1.59	2.41E-07
ILMN_2758070	Rtn1	Reticulon 1	1.57	6.48E-07
ILMN_2640862	Lrig1	Leucine-rich repeats and immunoglobulin-like	1.57	1.09E-05

		domains 1		
ILMN_2807665	Gatc	Glutamyl-tRNA(Gln) amidotransferase, subunit C homolog	1.57	1.38E-04
ILMN_1242617	Lpar4	Lysophosphatidic acid receptor 4	1.57	1.81E-04
ILMN_1244123	Slc38a2	Solute carrier family 38, member 2	1.57	9.30E-06
ILMN_2661588	Spnb3	spectrin beta 3	1.56	2.06E-04
ILMN_2552490	6720463L11Rik	Splicing factor, arginine/serine-rich 7 homolog	1.54	2.23E-05
ILMN_1226660	LOC100047402	Hypothetical protein LOC100047402	1.53	1.31E-04
ILMN_3115490	Rtn1	Reticulon 1	1.53	6.67E-06
ILMN_1226103	5730455P16Rik	Hypothetical protein LOC70591	1.53	1.72E-04

E9.5 mutant versus E8.5 wild type

Up-regulated transcripts (total = 41) ($p \leq 0.05$, fold change ≥ 1.5)

Illumina ID	Gene symbol	Name	Fold increase	P-value
ILMN_2500276	Ccng1	Cyclin G1	7.48	3.30E-05
ILMN_2634083	Cdkn1a	Cyclin dependent kinase inhibitor 1	6.87	3.21E-06
ILMN_2710229	Ccng1	Cyclin G1	5.74	6.01E-04
ILMN_2923607	Phlda3	Pleckstrin homology-like domain, family A, member 3	4.81	3.30E-05
ILMN_2702233	Ccng1	Cyclin G1	4.35	1.31E-03
ILMN_2846775	Cdkn1a	Cyclin dependent kinase inhibitor 1	4.00	3.43E-06
ILMN_2506012	Trp53inp1	transformation related protein 53 inducible nuclear protein 1	3.95	6.01E-04
ILMN_2628567	Phlda3	Pleckstrin homology-like domain, family A, member 3	3.71	3.41E-02
ILMN_2617256	Ms4a10	Membrane-spanning 4-domains, subfamily A, member 10	3.55	7.47E-03
ILMN_2846776	Cdkn1a	Cyclin dependent kinase inhibitor 1	3.25	3.18E-04
ILMN_2765454	Svop	Synaptic vesicle 2 related protein	3.22	1.90E-02
ILMN_2764551	Cox6b2	Cytochrome c oxidase subunit VIb polypeptide 2	3.01	6.89E-04
ILMN_2695819	Ddit4l	DNA-damage-inducible transcript 4-like	2.95	2.37E-04
ILMN_1250531	Slc19a2	Solute carrier family 19 (thiamine transporter), member 2	2.92	6.01E-04

ILMN_1217442	LOC674611	Similar to zinc finger protein 365	2.33	3.41E-02
ILMN_3163053	Commd3	COMM domain containing 3	2.30	2.82E-04
ILMN_2644936	Habp2	Hyaluronic acid binding protein 2	2.22	1.90E-02
ILMN_2690575	Pcbd1	Pterin 4 alpha carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1) 1	2.14	3.39E-02
ILMN_2647313	Klhl26	Kelch-like 26 (Drosophila)	2.12	1.98E-02
ILMN_2418426	Zmat3	Zinc finger matrin type 3	2.11	6.44E-03
ILMN_1247811	Ass1	Argininosuccinate synthetase 1	2.10	3.11E-03
ILMN_1226293	Slc39a4	Solute carrier family 39 (zinc transporter), member 4	2.06	3.41E-02
ILMN_2920736	Bbc3	BCL2 binding component 3	2.05	3.11E-03
ILMN_2979190	Habp2	Hyaluronic acid binding protein 2	2.04	6.93E-04
ILMN_2477136	2010003K15Rik	Hypothetical protein LOC75606	1.92	8.80E-04
ILMN_1245909	4933402E13Rik	Hypothetical protein LOC74437	1.83	4.78E-02
ILMN_2715546	Gpx3	Glutathione peroxidase 3	1.82	2.95E-02
ILMN_1246289	Got1l1	Glutamic-oxaloacetic transaminase 1-like 1	1.80	6.44E-03
ILMN_2619107	Lgals1	Lectin, galactose binding, soluble 1	1.76	3.88E-02
ILMN_2907655	Bax	BCL2-associated X protein	1.74	3.39E-02
ILMN_2908070	Gtse1	G two S phase expressed protein 1	1.73	6.53E-03
ILMN_2702102	D630023F18Rik	Hypothetical protein LOC98303	1.70	2.39E-03
ILMN_2604310	1810022C23Rik	Hypothetical protein LOC69123	1.69	4.91E-02
ILMN_2915833	BC025446	BC025446	1.67	4.78E-02
ILMN_2984462	Ercc5	Excision repair cross-complementing rodent repair deficiency, complementation group 5	1.64	1.68E-02
ILMN_2595842	Def6	Differentially expressed in FDCP 6	1.62	4.90E-02
ILMN_2604179	Lrrc3	Leucine rich repeat containing 3	1.61	4.09E-02
ILMN_2641371	Lgals2	Lectin, galactose-binding, soluble 2	1.61	2.25E-02
ILMN_1253618	Ccs	Copper chaperone for superoxide dismutase	1.60	2.21E-02
ILMN_2631093	Cd81	CD81 antigen	1.56	2.28E-02
ILMN_1255479	Cldn2	Claudin 2	1.53	3.44E-02

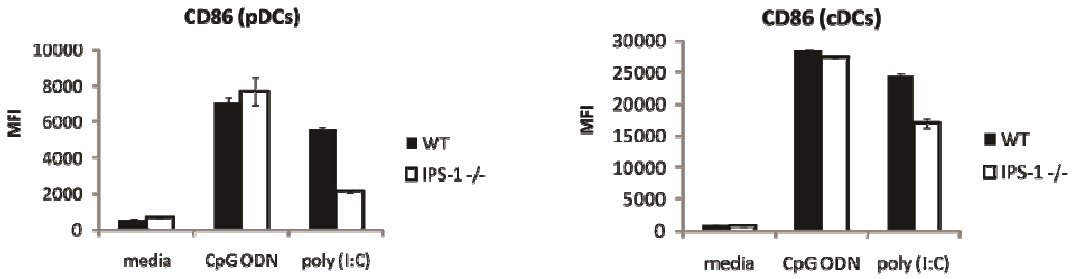
E9.5 mutant versus E8.5 wild type

Down-regulated transcripts (total = 9) ($p \leq 0.05$, fold change ≥ 1.5)

Illumina ID	Gene symbol	Name	Fold decrease	P value
ILMN_2622209	Rnaseh2b	Ribonuclease H2, subunit B	3.78	9.01E-07
ILMN_2622208	Rnaseh2b	Ribonuclease H2, subunit B	3.52	3.25E-07
ILMN_2812286	Fgf17	fibroblast growth factor 1	1.84	3.43E-05
ILMN_1220516	Wnt3a	wingless-related MMTV integration site 3A	1.83	1.59E-05
ILMN_2584527	Mbd2	Methyl-CpG binding domain protein 2	1.77	1.73E-05
ILMN_1233535	6330407J23Rik	Hypothetical protein LOC67412	1.75	8.38E-06
ILMN_1216132	Kpnb3	karyopherin (importin) beta 3	1.59	3.33E-05
ILMN_2640862	Lrig1	leucine-rich repeats and immunoglobulin-like domains 1	1.56	6.28E-05
ILMN_2776909	Irx3	Iroquois related homeobox 3	1.51	4.48E-06

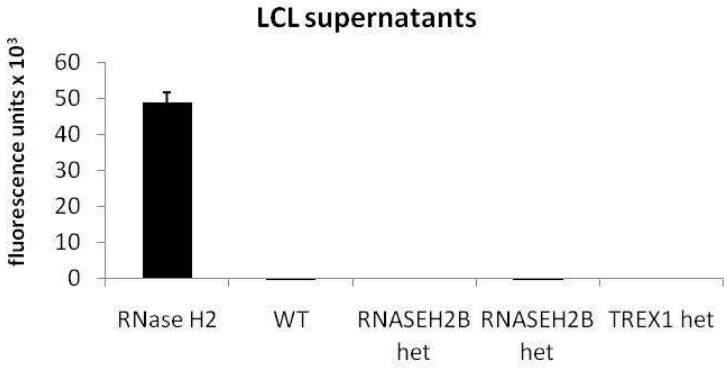
Appendix 11. Decreased expression of CD86 by IPS-1$^{-/-}$ FL-DCs in response to transfected poly (I:C)

As demonstrated in Figures 6.21 and 6.22, a range of PRR ligands were utilised in FL-DCs from IPS-1$^{-/-}$ mice but the cytokine response and phenotypic activation was equivalent to wild type. The only difference was observed in CD86 expression which was decreased in the IPS-1$^{-/-}$ cells compared to wild type. Data shown are from one experiment performed in duplicate $\pm$ s.d.



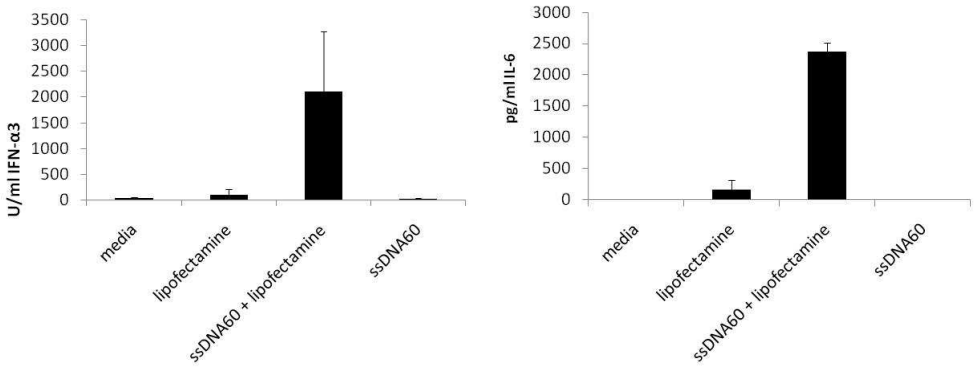
Appendix 12. RNase H2 activity is likely to be intracellular

The supernatants from EBV-transformed LCL lines were assayed for secreted RNase H2 activity using the RNase H activity assay. 5 μg of supernatant (as determined by protein concentration) were incubated with 25 pmol RNase H2-specific substrate and the fluorescence units released determined after 30 min incubation at 37^{o}C. 5 μg recombinant human RNase H2 was included as a control. Background fluorescence values were subtracted from each value. All supernatants were negative for RNase H2 activity. Data are from one experiment $\pm$ s.d. of triplicate wells. No activity was detected when 100 μg of supernatant was tested either, however these data only show that transformed LCLs do not secrete RNase H2 activity, and therefore cannot exclude the possibility that other cell types do.



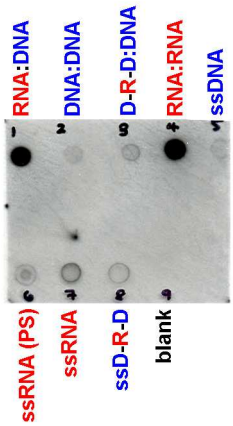
Appendix 13. ssDNA60 requires transfection to stimulate cytokine production by FL-DCs

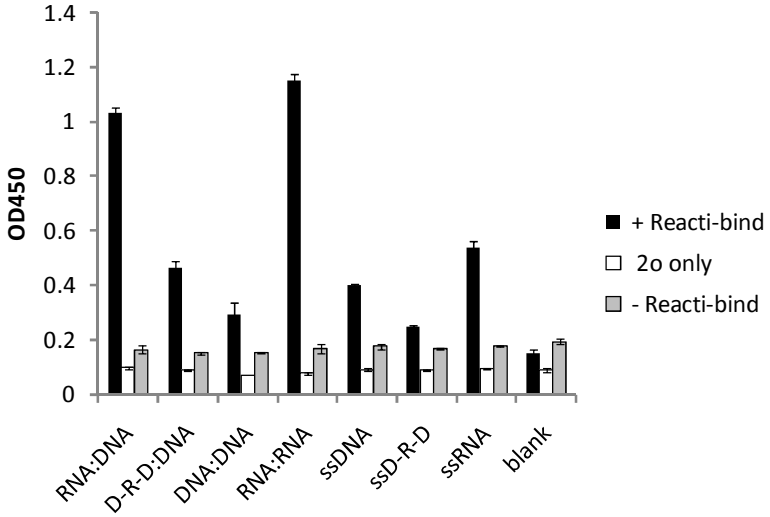
FL-DCs were stimulated with 1 μg/ml ssDNA60 in the presence or absence of Lipofectamine 2000 for 18 h. Cytokine production was measured by ELISA. Data are from one experiment performed in duplicate ± s.d.



Appendix 14. The detection of autoantibodies capable of recognising RNA:DNA hybrids in an SLE patient

The sera of three patients with SLE were tested for the presence of anti-nucleic acid autoantibodies using the dot blot and plate assay methods established in Chapter 4. High levels of autoantibodies against all nucleic-acid species was detected in the sera from one patient, used here at a dilution of 1:10,000. In particular, high levels of autoantibodies detected the dsRNA and RNA:DNA hybrid species. Data shown are representative of two independent experiments performed using sera from this patient. Plate assay data are presented as mean OD450 reading ± s.d. of triplicate wells.





8. Bibliography

Abbotts J., Bebenek K., Kunkel T. A. and Wilson S. H. (1993) Mechanism of HIV-1 reverse transcriptase. Termination of processive synthesis on a natural DNA template is influenced by the sequence of the template-primer stem. *J Biol Chem* 268, 10312-23.

Ablasser A., Bauernfeind F., Hartmann G., Latz E., Fitzgerald K. A. and Hornung V. (2009) RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nat Immunol* 10, 1065-1072.

Agrawal A., Eastman Q. M. and Schatz D. G. (1998) Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* 394, 744-51.

Ahn S. J., Costa J. and Emanuel J. R. (1996) PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. *Nucleic Acids Res* 24, 2623-5.

Aicardi J. (2002) Aicardi-Goutieres syndrome: special type early-onset encephalopathy. *Eur.J.Paediatr.Neurol.* 6 Suppl A, A1-A7.

Aicardi J. and Goutieres F. (1984) A progressive familial encephalopathy in infancy with calcifications of the basal ganglia and chronic cerebrospinal fluid lymphocytosis. *Ann.Neurol.* 15, 49-54.

Akwa Y., Hassett D. E., Eloranta M. L., Sandberg K., Masliah E., Powell H., Whitton J. L., Bloom F. E. and Campbell I. L. (1998) Transgenic expression of IFN-alpha in the central nervous system of mice protects against lethal neurotropic viral infection but induces inflammation and neurodegeneration. *J.Immunol.* 161, 5016-5026.

Al-Ahmadi W., Al-Haj L., Al-Mohanna F. A., Silverman R. H. and Khabar K. S. (2009) RNase L downmodulation of the RNA-binding protein, HuR, and cellular growth. *Oncogene* 28, 1782-91.

Alexopoulou L., Holt A. C., Medzhitov R. and Flavell R. A. (2001) Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* 413, 732-8.

Ali M., Highet L. J., Lacombe D., Goizet C., King M. D., Tacke U., van der Knaap M. S., Lagae L., Rittey C., Brunner H. G., van Bokhoven H., Hamel B., Oade Y. A., Sanchis A., Desguerre I., Cau D., Mathieu N., Moutard M. L., Lebon P., Kumar D., Jackson A. P. and Crow Y. J. (2006) A second locus for Aicardi-Goutieres syndrome at chromosome 13q14-21. *J.Med.Genet.* 43, 444-450.

Allen I. C., Scull M. A., Moore C. B., Holl E. K., McElvania-TeKippe E., Taxman D. J., Guthrie E. H., Pickles R. J. and Ting J. P. (2009) The NLRP3 inflammasome mediates in vivo innate immunity to influenza A virus through recognition of viral RNA. *Immunity* 30, 556-65.

Almofti M. R., Harashima H., Shinohara Y., Almofti A., Baba Y. and Kiwada H. (2003) Cationic liposome-mediated gene delivery: biophysical study and mechanism of internalization. *Arch Biochem Biophys* 410, 246-53.

Altschul S. F., Gish W., Miller W., Myers E. W. and Lipman D. J. (1990) Basic local alignment search tool. *J Mol Biol* 215, 403-10.

Andrejeva J., Childs K. S., Young D. F., Carlos T. S., Stock N., Goodbourn S. and Randall R. E. (2004) The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN-beta promoter. *Proc Natl Acad Sci U S A* 101, 17264-9.

Andrews-Pfannkoch C., Fadrosh D. W., Thorpe J. and Williamson S. J. (2010) Hydroxyapatite-mediated separation of double-stranded DNA, single-stranded DNA, and RNA genomes from natural viral assemblages. *Appl Environ Microbiol* 76, 5039-45.

Applequist S. E., Wallin R. P. and Ljunggren H. G. (2002) Variable expression of Toll-like receptor in murine innate and adaptive immune cell lines. *Int Immunol* 14, 1065-74.

Aravind L. and Koonin E. V. (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem Sci* 23, 469-72.

Arezi B. and Kuchta R. D. (2000) Eukaryotic DNA primase. *Trends Biochem Sci* 25, 572-6.

Arnott S., Chandrasekaran R., Millane R. P. and Park H. S. (1986) RNA-RNA, DNA-DNA, and DNA-RNA Polymorphism. *Biophys J* 49, 3-5.

Arudchandran A., Cerritelli S., Narimatsu S., Itaya M., Shin D. Y., Shimada Y. and Crouch R. J. (2000) The absence of ribonuclease H1 or H2 alters the sensitivity of Saccharomyces cerevisiae to hydroxyurea, caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. *Genes Cells* 5, 789-802.

Ashburner M., Ball C. A., Blake J. A., Botstein D., Butler H., Cherry J. M., Davis A. P., Dolinski K., Dwight S. S., Eppig J. T., Harris M. A., Hill D. P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J. C., Richardson J. E., Ringwald M., Rubin G. M. and Sherlock G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25, 25-9.

Asselin-Paturel C., Boonstra A., Dalod M., Durand I., Yessaad N., Dezutter-Dambuyant C., Vicari A., O'Garra A., Biron C., Briere F. and Trinchieri G. (2001) Mouse type I IFN-producing cells are immature APCs with plasmacytoid morphology. *Nat Immunol* 2, 1144-50.

Ayyagari R., Gomes X. V., Gordenin D. A. and Burgers P. M. (2003) Okazaki fragment maturation in yeast. I. Distribution of functions between FEN1 AND DNA2. *J Biol Chem* 278, 1618-25.

Azzalin C. M., Reichenbach P., Khoriauli L., Giulotto E. and Lingner J. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 318, 798-801.

Bachmeyer C., Farge D., Gluckman E., Miclea J. M. and Aractingi S. (1996) Raynaud's phenomenon and digital necrosis induced by interferon-alpha. *Br.J.Dermatol.* 135, 481-483.

Bae S. H., Bae K. H., Kim J. A. and Seo Y. S. (2001) RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. *Nature* 412, 456-61.

Bae S. H. and Seo Y. S. (2000) Characterization of the enzymatic properties of the yeast dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing. *J Biol Chem* 275, 38022-31.

Baker K. P., Baron W. F., Henzel W. J. and Spencer S. A. (1998) Molecular cloning and characterization of human and murine DNase II. *Gene* 215, 281-9.

Baker T. A. and Kornberg A. (1988) Transcriptional activation of initiation of replication from the E. coli chromosomal origin: an RNA-DNA hybrid near oriC. *Cell* 55, 113-23.

Balada E., Ordi-Ros J. and Vilardell-Tarres M. (2009) Molecular mechanisms mediated by human endogenous retroviruses (HERVs) in autoimmunity. *Rev Med Virol* 19, 273-286.

Bale J. F., Jr., Bray P. F. and Bell W. E. (1985) Neuroradiographic abnormalities in congenital cytomegalovirus infection. *Pediatr Neurol* 1, 42-7.

Bambara R. A., Murante R. S. and Henricksen L. A. (1997) Enzymes and reactions at the eukaryotic DNA replication fork. *J.Biol.Chem.* 272, 4647-4650.

Bannert N. and Kurth R. (2004) Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A* 101 Suppl 2, 14572-9.

Barrat F. J., Meeker T., Gregorio J., Chan J. H., Uematsu S., Akira S., Chang B., Duramad O. and Coffman R. L. (2005) Nucleic acids of mammalian origin can act as endogenous ligands for Toll-like receptors and may promote systemic lupus erythematosus. *J Exp Med* 202, 1131-9.

Barton G. M., Kagan J. C. and Medzhitov R. (2006) Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nat Immunol* 7, 49-56.

Bauer S., Kirschning C. J., Hacker H., Redecke V., Hausmann S., Akira S., Wagner H. and Lipford G. B. (2001) Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. *Proc Natl Acad Sci U S A* 98, 9237-42.

Baust C., Gagnier L., Baillie G. J., Harris M. J., Juriloff D. M. and Mager D. L. (2003) Structure and expression of mobile ETnII retroelements and their coding-competent MusD relatives in the mouse. *J Virol* 77, 11448-58.

Beignon A. S., McKenna K., Skoberne M., Manches O., DaSilva I., Kavanagh D. G., Larsson M., Gorelick R. J., Lifson J. D. and Bhardwaj N. (2005) Endocytosis of HIV-1 activates plasmacytoid dendritic cells via Toll-like receptor-viral RNA interactions. *J Clin Invest* 115, 3265-75.

Belanger K. G. and Kreuzer K. N. (1998) Bacteriophage T4 initiates bidirectional DNA replication through a two-step process. *Mol Cell* 2, 693-701.

Beld M., Sol C., Goudsmit J. and Boom R. (1996) Fractionation of nucleic acids into single-stranded and double-stranded forms. *Nucleic Acids Res* 24, 2618-9.

Belman A. L., Lantos G., Horoupian D., Novick B. E., Ultmann M. H., Dickson D. W. and Rubinstein A. (1986) AIDS: calcification of the basal ganglia in infants and children. *Neurology* 36, 1192-9.

Belman A. L., Muenz L. R., Marcus J. C., Goedert J. J., Landesman S., Rubinstein A., Goodwin S., Durako S. and Willoughby A. (1996) Neurologic status of human immunodeficiency virus 1-infected infants and their controls: a prospective study from birth to 2 years. Mothers and Infants Cohort Study. *Pediatrics* 98, 1109-18.

Benson D. A., Karsch-Mizrachi I., Lipman D. J., Ostell J. and Wheeler D. L. (2008) GenBank. *Nucleic Acids Res* 36, D25-30.
Bergoglio V., Ferrari E., Hubscher U., Cazaux C. and Hoffmann J. S. (2003) DNA polymerase beta can incorporate ribonucleotides during DNA synthesis of undamaged and CPD-damaged DNA. *J Mol Biol* 331, 1017-23.
Bernad A., Blanco L., Lazaro J. M., Martin G. and Salas M. (1989) A conserved 3'----5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* 59, 219-28.
Bernardi G. (1965) Chromatography of nucleic acids on hydroxyapatite. *Nature* 206, 779-83.
Bernardi G. (1968) Mechanism of action and structure of acid deoxyribonuclease. *Adv Enzymol Relat Areas Mol Biol* 31, 1-49.
Bernstein E., Caudy A. A., Hammond S. M. and Hannon G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6.
Bhoj V. G. and Chen Z. J. (2008) Linking retroelements to autoimmunity. *Cell* 134, 569-71.
Bisbal C., Silhol M., Laubenthal H., Kaluza T., Carnac G., Milligan L., Le Roy F. and Salehzada T. (2000) The 2'-5' oligoadenylate/RNase L/RNase L inhibitor pathway regulates both MyoD mRNA stability and muscle cell differentiation. *Mol Cell Biol* 20, 4959-69.
Blanco L., Bernad A. and Salas M. (1992) Evidence favouring the hypothesis of a conserved 3'-5' exonuclease active site in DNA-dependent DNA polymerases. *Gene* 112, 139-44.
Blasius A. L. and Beutler B. (2010) Intracellular toll-like receptors. *Immunity* 32, 305-15.
Blau N., Bonafe L., Krageloh-Mann I., Thony B., Kierat L., Hausler M. and Ramaekers V. (2003) Cerebrospinal fluid pterins and folates in Aicardi-Goutieres syndrome: a new phenotype. *Neurology* 61, 642-7.
Boczkowski D., Nair S. K., Snyder D. and Gilboa E. (1996) Dendritic cells pulsed with RNA are potent antigen-presenting cells in vitro and in vivo. *J Exp Med* 184, 465-72.
Bodano A., Gonzalez A., Ferreiros-Vidal I., Balada E., Ordi J., Carreira P., Gomez-Reino J. J. and Conde C. (2006) Association of a non-synonymous single-nucleotide polymorphism of DNASEI with SLE susceptibility. *Rheumatology (Oxford)* 45, 819-23.
Boeke J. D. (2003) The unusual phylogenetic distribution of retrotransposons: a hypothesis. *Genome Res* 13, 1975-83.
Bogdan C. (2000) The function of type I interferons in antimicrobial immunity. *Curr Opin Immunol* 12, 419-24.
Bogdan C., Mattner J. and Schleicher U. (2004) The role of type I interferons in non-viral infections. *Immunol Rev* 202, 33-48.
Boguslawski S. J., Smith D. E., Michalak M. A., Mickelson K. E., Yehle C. O., Patterson W. L. and Carrico R. J. (1986) Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J Immunol Methods* 89, 123-30.
Boom R., Sol C. J., Salimans M. M., Jansen C. L., Wertheim-van Dillen P. M. and van der Noordaa J. (1990) Rapid and simple method for purification of nucleic acids. *J Clin Microbiol* 28, 495-503.
Borodina T. A., Lehrach H. and Soldatov A. V. (2003) DNA purification on homemade silica spin-columns. *Anal Biochem* 321, 135-7.
Boule J. B. and Zakian V. A. (2007) The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. *Nucleic Acids Res* 35, 5809-18.
Bourke E., Bosisio D., Golay J., Polentarutti N. and Mantovani A. (2003) The toll-like receptor repertoire of human B lymphocytes: inducible and selective expression of TLR9 and TLR10 in normal and transformed cells. *Blood* 102, 956-63.
Boztug K., Carson M. J., Pham-Mitchell N., Asensio V. C., DeMartino J. and Campbell I. L. (2002) Leukocyte infiltration, but not neurodegeneration, in the CNS of transgenic mice with astrocyte production of the CXC chemokine ligand 10. *J Immunol* 169, 1505-15.
Brasel K., De Smedt T., Smith J. L. and Maliszewski C. R. (2000) Generation of murine dendritic cells from flt3-ligand-supplemented bone marrow cultures. *Blood* 96, 3029-39.
Brawand P., Fitzpatrick D. R., Greenfield B. W., Brasel K., Maliszewski C. R. and De Smedt T. (2002) Murine plasmacytoid pre-dendritic cells generated from Flt3 ligand-supplemented bone marrow cultures are immature APCs. *J Immunol* 169, 6711-9.
Brinkmann M. M., Spooner E., Hoebe K., Beutler B., Ploegh H. L. and Kim Y. M. (2007) The interaction between the ER membrane protein UNC93B and TLR3, 7, and 9 is crucial for TLR signaling. *J Cell Biol* 177, 265-75.

Broccoli S., Rallu F., Sanscartier P., Cerritelli S. M., Crouch R. J. and Drolet M. (2004) Effects of RNA polymerase modifications on transcription-induced negative supercoiling and associated R-loop formation. *Mol.Microbiol.* 52, 1769-1779.

Brosh R. M., Jr., von Kobbe C., Sommers J. A., Karmakar P., Opresko P. L., Piotrowski J., Dianova I., Dianov G. L. and Bohr V. A. (2001) Werner syndrome protein interacts with human flap endonuclease 1 and stimulates its cleavage activity. *EMBO J* 20, 5791-801.

Brown T. A., Tkachuk A. N. and Clayton D. A. (2008) Native R-loops persist throughout the mouse mitochondrial DNA genome. *J Biol Chem* 283, 36743-36751.

Brucet M., Querol-Audi J., Serra M., Ramirez-Espain X., Bertlik K., Ruiz L., Lloberas J., Macias M. J., Fita I. and Celada A. (2007) Structure of the dimeric exonuclease TREX1 in complex with DNA displays a proline-rich binding site for WW domains. *J Biol Chem*.

Bsibsi M., Ravid R., Gveric D. and van Noort J. M. (2002) Broad expression of Toll-like receptors in the human central nervous system. *J Neuropathol Exp Neurol* 61, 1013-21.

Burckstummer T., Baumann C., Bluml S., Dixit E., Durnberger G., Jahn H., Planyavsky M., Bilban M., Colinge J., Bennett K. L. and Superti-Furga G. (2009) An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nat Immunol* 10, 266-72.

Burgers P. M. (2009) Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem* 284, 4041-5.

Buscher K., Trefzer U., Hofmann M., Sterry W., Kurth R. and Denner J. (2005) Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. *Cancer Res* 65, 4172-80.

Campbell I. L., Krucker T., Steffensen S., Akwa Y., Powell H. C., Lane T., Carr D. J., Gold L. H., Henriksen S. J. and Siggins G. R. (1999) Structural and functional neuropathology in transgenic mice with CNS expression of IFN-alpha. *Brain Res.* 835, 46-61.

Campo-Voegeli A., Estrach T., Marti R. M., Corominas N., Tuset M. and Mascaro J. M. (1998) Acrocyanosis induced by interferon alpha(2a). *Dermatology* 196, 361-363.

Cao X. (2009) New DNA-sensing pathway feeds RIG-I with RNA. *Nat Immunol* 10, 1049-51.

Carpentier P. A., Begolka W. S., Olson J. K., Elhofy A., Karpus W. J. and Miller S. D. (2005) Differential activation of astrocytes by innate and adaptive immune stimuli. *Glia* 49, 360-74.

Carpten J., Nupponen N., Isaacs S., Sood R., Robbins C., Xu J., Faruque M., Moses T., Ewing C., Gillanders E., Hu P., Bujnovszky P., Makalowska I., Baffoe-Bonnie A., Faith D., Smith J., Stephan D., Wiley K., Brownstein M., Gildea D., Kelly B., Jenkins R., Hostetter G., Matikainen M., Schleutker J., Klinger K., Connors T., Xiang Y., Wang Z., De Marzo A., Papadopoulos N., Kallioniemi O. P., Burk R., Meyers D., Gronberg H., Meltzer P., Silverman R., Bailey-Wilson J., Walsh P., Isaacs W. and Trent J. (2002) Germline mutations in the ribonuclease L gene in families showing linkage with HPC1. *Nat Genet* 30, 181-4.

Casey G., Neville P. J., Plummer S. J., Xiang Y., Krumroy L. M., Klein E. A., Catalona W. J., Nupponen N., Carpten J. D., Trent J. M., Silverman R. H. and Witte J. S. (2002) RNASEL Arg462Gln variant is implicated in up to 13% of prostate cancer cases. *Nat Genet* 32, 581-3.

Cate J. H., Gooding A. R., Podell E., Zhou K., Golden B. L., Kundrot C. E., Cech T. R. and Doudna J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* 273, 1678-85.

Cazenave C. and Uhlenbeck O. C. (1994) RNA template-directed RNA synthesis by T7 RNA polymerase. *Proc Natl Acad Sci U S A* 91, 6972-6.

Cella M., Jarrossay D., Facchetti F., Alebardi O., Nakajima H., Lanzavecchia A. and Colonna M. (1999) Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon. *Nat Med* 5, 919-23.

Cerritelli S. M. and Crouch R. J. (1998) Cloning, expression, and mapping of ribonucleases H of human and mouse related to bacterial RNase HI. *Genomics* 53, 300-307.

Cerritelli S. M. and Crouch R. J. (2009) Ribonuclease H: the enzymes in eukaryotes. *Febs J* 276, 1494-505.

Cerritelli S. M., Frolova E. G., Feng C., Grinberg A., Love P. E. and Crouch R. J. (2003) Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell* 11, 807-15.

Chaiet L. and Wolf F. J. (1964) The Properties of Streptavidin, a Biotin-Binding Protein Produced by Streptomycetes. *Arch Biochem Biophys* 106, 1-5.

Chakraborty P. and Grosse F. (2010) WRN helicase unwinds Okazaki fragment-like hybrids in a reaction stimulated by the human DHX9 helicase. *Nucleic Acids Res*.

Chan O. T., Madaio M. P. and Shlomchik M. J. (1999) The central and multiple roles of B cells in lupus pathogenesis. *Immunol Rev* 169, 107-21.
Chapados B. R., Chai Q., Hosfield D. J., Qiu J., Shen B. and Tainer J. A. (2001) Structural biochemistry of a type 2 RNase H: RNA primer recognition and removal during DNA replication. *J.Mol.Biol.* 307, 541-556.
Cherkasova V., Lyons D. M. and Elion E. A. (1999) Fus3p and Kss1p control G1 arrest in Saccharomyces cerevisiae through a balance of distinct arrest and proliferative functions that operate in parallel with Far1p. *Genetics* 151, 989-1004.
Chitrabamrung S., Rubin R. L. and Tan E. M. (1981) Serum deoxyribonuclease I and clinical activity in systemic lupus erythematosus. *Rheumatol Int* 1, 55-60.
Chiu Y. H., Macmillan J. B. and Chen Z. J. (2009) RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway. *Cell* 138, 576-91.
Choi M. K., Wang Z., Ban T., Yanai H., Lu Y., Koshiba R., Nakaima Y., Hangai S., Savitsky D., Nakasato M., Negishi H., Takeuchi O., Honda K., Akira S., Tamura T. and Taniguchi T. (2009) A selective contribution of the RIG-I-like receptor pathway to type I interferon responses activated by cytosolic DNA. *Proc Natl Acad Sci U S A* 106, 17870-5.
Chon H., Vassilev A., DePamphilis M. L., Zhao Y., Zhang J., Burgers P. M., Crouch R. J. and Cerritelli S. M. (2009) Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex. *Nucleic Acids Res* 37, 96-110.
Chou S. H., Flynn P. and Reid B. (1989) High-resolution NMR study of a synthetic DNA-RNA hybrid dodecamer containing the consensus pribnow promoter sequence: d(CGTTATAATGCG).r(CGCAUUAUAACG). *Biochemistry* 28, 2435-43.
Choudhary C., Kumar C., Gnad F., Nielsen M. L., Rehman M., Walther T. C., Olsen J. V. and Mann M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834-40.
Chowdhury D., Beresford P. J., Zhu P., Zhang D., Sung J. S., Demple B., Perrino F. W. and Lieberman J. (2006) The exonuclease TREX1 is in the SET complex and acts in concert with NM23-H1 to degrade DNA during granzyme A-mediated cell death. *Mol Cell* 23, 133-42.
Chowdhury D. and Lieberman J. (2008) Death by a thousand cuts: granzyme pathways of programmed cell death. *Annu Rev Immunol* 26, 389-420.
Christensen S. R., Shupe J., Nickerson K., Kashgarian M., Flavell R. A. and Shlomchik M. J. (2006) Toll-like receptor 7 and TLR9 dictate autoantibody specificity and have opposing inflammatory and regulatory roles in a murine model of lupus. *Immunity* 25, 417-28.
Christmann M., Tomicic M. T., Aasland D., Berdelle N. and Kaina B. (2010) Three prime exonuclease I (TREX1) is Fos/AP-1 regulated by genotoxic stress and protects against ultraviolet light and benzo(a)pyrene-induced DNA damage. *Nucleic Acids Res* 38, 6418-32.
Chung A. and Older S. A. (1997) Interferon-alpha associated arthritis. *J Rheumatol* 24, 1844-5.
Coccia E. M., Severa M., Giacomini E., Monneron D., Remoli M. E., Julkunen I., Cella M., Lande R. and Uze G. (2004) Viral infection and Toll-like receptor agonists induce a differential expression of type I and lambda interferons in human plasmacytoid and monocyte-derived dendritic cells. *Eur J Immunol* 34, 796-805.
Colby C., Stollar B. D. and Simon M. I. (1971) Interferon induction: DNA-RNA hybrid or double stranded RNA? *Nat New Biol* 229, 172-4.
Cole J. L. (2007) Activation of PKR: an open and shut case? *Trends Biochem Sci* 32, 57-62.
Coonrod A., Li F. Q. and Horwitz M. (1997) On the mechanism of DNA transfection: efficient gene transfer without viruses. *Gene Ther* 4, 1313-21.
Coutlee F., Bobo L., Abbass H., Dalabetta G., Hook N. E., Shah K. and Viscidi R. P. (1992) Detection of HPV-16 in cell lines and cervical lavage specimens by a polymerase chain reaction-enzyme immunoassay assay. *J Med Virol* 37, 22-9.
Coutlee F., Bobo L., Mayur K., Yolken R. H. and Viscidi R. P. (1989) Immunodetection of DNA with biotinylated RNA probes: a study of reactivity of a monoclonal antibody to DNA-RNA hybrids. *Anal Biochem* 181, 96-105.
Coutlee F., Yang B. Z., Bobo L., Mayur K., Yolken R. and Viscidi R. (1990) Enzyme immunoassay for detection of hybrids between PCR-amplified HIV-1 DNA and a RNA probe: PCR-EIA. *AIDS Res Hum Retroviruses* 6, 775-84.

Crow Y. J., Hayward B. E., Parmar R., Robins P., Leitch A., Ali M., Black D. N., van Bokhoven H., Brunner H. G., Hamel B. C., Corry P. C., Cowan F. M., Frints S. G., Klepper J., Livingston J. H., Lynch S. A., Massey R. F., Meritet J. F., Michaud J. L., Ponsot G., Voit T., Lebon P., Bonthron D. T., Jackson A. P., Barnes D. E. and Lindahl T. (2006a) Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus. *Nat.Genet.* 38, 917-920.

Crow Y. J., Jackson A. P., Roberts E., van Beusekom E., Barth P., Corry P., Ferrie C. D., Hamel B. C., Jayatunga R., Karbani G., Kalmanchey R., Kelemen A., King M., Kumar R., Livingstone J., Massey R., McWilliam R., Meager A., Rittey C., Stephenson J. B., Tolmie J. L., Verrips A., Voit T., van Bokhoven H., Brunner H. G. and Woods C. G. (2000) Aicardi-Goutieres syndrome displays genetic heterogeneity with one locus (AGS1) on chromosome 3p21. *Am.J.Hum.Genet.* 67, 213-221.

Crow Y. J., Leitch A., Hayward B. E., Garner A., Parmar R., Griffith E., Ali M., Semple C., Aicardi J., Babul-Hirji R., Baumann C., Baxter P., Bertini E., Chandler K. E., Chitayat D., Cau D., Dery C., Fazzi E., Goizet C., King M. D., Klepper J., Lacombe D., Lanzi G., Lyall H., Martinez-Frias M. L., Mathieu M., McKeown C., Monier A., Oade Y., Quarrell O. W., Rittey C. D., Rogers R. C., Sanchis A., Stephenson J. B., Tacke U., Till M., Tolmie J. L., Tomlin P., Voit T., Weschke B., Woods C. G., Lebon P., Bonthron D. T., Ponting C. P. and Jackson A. P. (2006b) Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. *Nat.Genet.* 38, 910-916.

Crow Y. J. and Livingston J. H. (2008) Aicardi-Goutieres syndrome: an important Mendelian mimic of congenital infection. *Dev Med Child Neurol* 50, 410-6.

Crow Y. J. and Rehwinkel J. (2009) Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. *Hum Mol Genet* 18, R130-6.

Cui S., Eisenacher K., Kirchhofer A., Brzozka K., Lammens A., Lammens K., Fujita T., Conzelmann K. K., Krug A. and Hopfner K. P. (2008) The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I. *Mol Cell* 29, 169-79.

Cutrone E. C. and Langer J. A. (1997) Contributions of cloned type I interferon receptor subunits to differential ligand binding. *FEBS Lett* 404, 197-202.

D'Cruz D. P., Khamashta M. A. and Hughes G. R. (2007) Systemic lupus erythematosus. *Lancet* 369, 587-96.

Dale R. C., Brilot F., Fagan E. and Earl J. (2009) Cerebrospinal fluid neopterin in paediatric neurology: a marker of active central nervous system inflammation. *Dev Med Child Neurol* 51, 317-23.

Daniels G. A. and Lieber M. R. (1995) RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. *Nucleic Acids Res* 23, 5006-11.

de Lange T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19, 2100-10.

de Parseval N., Lazar V., Casella J. F., Benit L. and Heidmann T. (2003) Survey of human genes of retroviral origin: identification and transcriptome of the genes with coding capacity for complete envelope proteins. *J Virol* 77, 10414-22.

de Silva U., Choudhury S., Bailey S. L., Harvey S., Perrino F. W. and Hollis T. (2007) The crystal structure of TREX1 explains the 3' nucleotide specificity and reveals a polyproline II helix for protein partnering. *J Biol Chem* 282, 10537-43.

de Veer M. J., Holko M., Frevel M., Walker E., Der S., Paranjape J. M., Silverman R. H. and Williams B. R. (2001) Functional classification of interferon-stimulated genes identified using microarrays. *J Leukoc Biol* 69, 912-20.

DeFilippis V. R., Alvarado D., Sali T., Rothenburg S. and Fruh K. (2010) Human cytomegalovirus induces the interferon response via the DNA sensor ZBP1. *J Virol* 84, 585-98.

Deininger P. L. and Batzer M. A. (2002) Mammalian retroelements. *Genome Res* 12, 1455-65.

Delaloye J., Roger T., Steiner-Tardivel Q. G., Le Roy D., Knaup Reymond M., Akira S., Petrilli V., Gomez C. E., Perdiguero B., Tschopp J., Pantaleo G., Esteban M. and Calandra T. (2009) Innate immune sensing of modified vaccinia virus Ankara (MVA) is mediated by TLR2-TLR6, MDA-5 and the NALP3 inflammasome. *PLoS Pathog* 5, e1000480.

Dennis G., Jr., Sherman B. T., Hosack D. A., Yang J., Gao W., Lane H. C. and Lempicki R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4, P3.

DeStefano J. J., Buiser R. G., Mallaber L. M., Fay P. J. and Bambara R. A. (1992) Parameters that influence processive synthesis and site-specific termination by human immunodeficiency virus reverse transcriptase on RNA and DNA templates. *Biochim Biophys Acta* 1131, 270-80.
Di Domizio J., Blum A., Gallagher-Gambarelli M., Molens J. P., Chaperot L. and Plumas J. (2009) TLR7 stimulation in human plasmacytoid dendritic cells leads to the induction of early IFN-inducible genes in the absence of type I IFN. *Blood* 114, 1794-802.
Diaz-Uriate R., Al-Shahrour F. and Dopazo J. (2003) *Use of GO Terms to Understand the Biological Significance of Microarray Differential Gene Expression Data*. Kluwer.
Diebold S. S., Kaisho T., Hemmi H., Akira S. and Reis e Sousa C. (2004) Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* 303, 1529-31.
Diebold S. S., Massacrier C., Akira S., Paturel C., Morel Y. and Reis e Sousa C. (2006) Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides. *Eur J Immunol* 36, 3256-67.
Diebold S. S., Montoya M., Unger H., Alexopoulou L., Roy P., Haswell L. E., Al-Shamkhani A., Flavell R., Borrow P. and Reis e Sousa C. (2003) Viral infection switches non-plasmacytoid dendritic cells into high interferon producers. *Nature* 424, 324-8.
Dittmar M., Bischofs C., Matheis N., Poppe R. and Kahaly G. J. (2008) A novel mutation in the DNASE1 gene is related with protein instability and decreased enzyme activity in thyroid autoimmunity. *J Autoimmun* 32, 7-13.
Dong B., Kim S., Hong S., Das Gupta J., Malathi K., Klein E. A., Ganem D., Derisi J. L., Chow S. A. and Silverman R. H. (2007) An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors. *Proc Natl Acad Sci U S A* 104, 1655-60.
Douville R. N. and Hiscott J. (2010) The interface between the innate interferon response and expression of host retroviral restriction factors. *Cytokine* 52, 108-15.
Drexler S. K. and Foxwell B. M. (2010) The role of toll-like receptors in chronic inflammation. *Int J Biochem Cell Biol* 42, 506-18.
Du X., Poltorak A., Wei Y. and Beutler B. (2000) Three novel mammalian toll-like receptors: gene structure, expression, and evolution. *Eur Cytokine Netw* 11, 362-71.
Dudas K. C. and Kreuzer K. N. (2001) UvsW protein regulates bacteriophage T4 origin-dependent replication by unwinding R-loops. *Mol Cell Biol* 21, 2706-15.
Dugan J. W., Albor A., David L., Fowlkes J., Blackledge M. T., Martin T. M., Planck S. R., Rosenzweig H. L., Rosenbaum J. T. and Davey M. P. (2009) Nucleotide oligomerization domain-2 interacts with 2'-5'-oligoadenylate synthetase type 2 and enhances RNase-L function in THP-1 cells. *Mol Immunol* 47, 560-6.
Dunning M. J., Smith M. L., Ritchie M. E. and Tavare S. (2007) beadarray: R classes and methods for Illumina bead-based data. *Bioinformatics* 23, 2183-4.
Dutrow N., Nix D. A., Holt D., Milash B., Dalley B., Westbroek E., Parnell T. J. and Cairns B. R. (2008) Dynamic transcriptome of Schizosaccharomyces pombe shown by RNA-DNA hybrid mapping. *Nat Genet* 40, 977-86.
Dyer K. D. and Rosenberg H. F. (2006) The RNase a superfamily: generation of diversity and innate host defense. *Mol Divers* 10, 585-97.
Eder P. S. and Walder J. A. (1991) Ribonuclease H from K562 human erythroleukemia cells. Purification, characterization, and substrate specificity. *J.Biol.Chem.* 266, 6472-6479.
Eder P. S., Walder R. Y. and Walder J. A. (1993) Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. *Biochimie* 75, 123-6.
Edwards A. D., Diebold S. S., Slack E. M., Tomizawa H., Hemmi H., Kaisho T., Akira S. and Reis e Sousa C. (2003) Toll-like receptor expression in murine DC subsets: lack of TLR7 expression by CD8 alpha+ DC correlates with unresponsiveness to imidazoquinolines. *Eur J Immunol* 33, 827-33.
Egli M., Portmann S. and Usman N. (1996) RNA hydration: a detailed look. *Biochemistry* 35, 8489-94.
Egli M., Usman N. and Rich A. (1993) Conformational influence of the ribose 2'-hydroxyl group: crystal structures of DNA-RNA chimeric duplexes. *Biochemistry* 32, 3221-37.
Ehlers M. and Ravetch J. V. (2007) Opposing effects of Toll-like receptor stimulation induce autoimmunity or tolerance. *Trends Immunol* 28, 74-9.
El Hage A., French S. L., Beyer A. L. and Tollervey D. (2010) Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev* 24, 1546-58.

Evans C. J. and Aguilera R. J. (2003) DNase II: genes, enzymes and function. *Gene* 322, 1-15.
Ewald S. E., Lee B. L., Lau L., Wickliffe K. E., Shi G. P., Chapman H. A. and Barton G. M. (2008) The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. *Nature*.
Fairhurst A. M., Hwang S. H., Wang A., Tian X. H., Boudreaux C., Zhou X. J., Casco J., Li Q. Z., Connolly J. E. and Wakeland E. K. (2008) Yaa autoimmune phenotypes are conferred by overexpression of TLR7. *Eur J Immunol* 38, 1971-8.
Fan H. (2007) A new human retrovirus associated with prostate cancer. *Proc Natl Acad Sci U S A* 104, 1449-50.
Fedoroff O., Salazar M. and Reid B. R. (1993) Structure of a DNA:RNA hybrid duplex. Why RNase H does not cleave pure RNA. *J Mol Biol* 233, 509-23.
Fedoroff O. Y., Ge Y. and Reid B. R. (1997) Solution structure of r(gaggacug):d(CAGTCCTC) hybrid: implications for the initiation of HIV-1 (+)-strand synthesis. *J Mol Biol* 269, 225-39.
Fernandes-Alnemri T., Yu J. W., Datta P., Wu J. and Alnemri E. S. (2009) AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* 458, 509-13.
Finn R. D., Mistry J., Tate J., Coggill P., Heger A., Pollington J. E., Gavin O. L., Gunasekaran P., Ceric G., Forslund K., Holm L., Sonnhammer E. L., Eddy S. R. and Bateman A. (2010) The Pfam protein families database. *Nucleic Acids Res* 38, D211-22.
Floyd-Smith G., Slattery E. and Lengyel P. (1981) Interferon action: RNA cleavage pattern of a (2'-5')oligoadenylate--dependent endonuclease. *Science* 212, 1030-2.
Forsbach A., Nemorin J. G., Montino C., Muller C., Samulowitz U., Vicari A. P., Jurk M., Mutwiri G. K., Krieg A. M., Lipford G. B. and Vollmer J. (2008) Identification of RNA sequence motifs stimulating sequence-specific TLR8-dependent immune responses. *J Immunol* 180, 3729-38.
Franchi L., Eigenbrod T., Munoz-Planillo R. and Nunez G. (2009) The inflammasome: a caspase-1-activation platform that regulates immune responses and disease pathogenesis. *Nat Immunol* 10, 241-7.
Frank P., Braunshofer-Reiter C., Poltl A. and Holzmann K. (1998a) Cloning, subcellular localization and functional expression of human RNase HII. *Biol.Chem.* 379, 1407-1412.
Frank P., Braunshofer-Reiter C. and Wintersberger U. (1998b) Yeast RNase H(35) is the counterpart of the mammalian RNase HI, and is evolutionarily related to prokaryotic RNase HII. *FEBS Lett.* 421, 23-26.
Frank P., Braunshofer-Reiter C., Wintersberger U., Grimm R. and Busen W. (1998c) Cloning of the cDNA encoding the large subunit of human RNase HI, a homologue of the prokaryotic RNase HII. *Proc Natl Acad Sci U S A* 95, 12872-7.
Fugmann S. D. and Schatz D. G. (2003) Rna Aids DNA. *Nat Immunol* 4, 429-30.
Fukui R., Saitoh S., Matsumoto F., Kozuka-Hata H., Oyama M., Tabeta K., Beutler B. and Miyake K. (2009) Unc93B1 biases Toll-like receptor responses to nucleic acid in dendritic cells toward DNA- but against RNA-sensing. *J Exp Med* 206, 1339-50.
Funami K., Matsumoto M., Oshiumi H., Akazawa T., Yamamoto A. and Seya T. (2004) The cytoplasmic 'linker region' in Toll-like receptor 3 controls receptor localization and signaling. *Int Immunol* 16, 1143-54.
Gaidamakov S. A., Gorshkova, II, Schuck P., Steinbach P. J., Yamada H., Crouch R. J. and Cerritelli S. M. (2005) Eukaryotic RNases H1 act processively by interactions through the duplex RNA-binding domain. *Nucleic Acids Res* 33, 2166-75.
Gaipl U. S., Beyer T. D., Heyder P., Kuenkele S., Bottcher A., Voll R. E., Kalden J. R. and Herrmann M. (2004) Cooperation between C1q and DNase I in the clearance of necrotic cell-derived chromatin. *Arthritis Rheum* 50, 640-9.
Ganguly D., Chamilos G., Lande R., Gregorio J., Meller S., Facchinetti V., Homey B., Barrat F. J., Zal T. and Gilliet M. (2009) Self-RNA-antimicrobial peptide complexes activate human dendritic cells through TLR7 and TLR8. *J Exp Med* 206, 1983-94.
Geijtenbeek T. B. (2010) Host DNase TREX1 hides HIV from DNA sensors. *Nat Immunol* 11, 979-80.
Genin P., Vaccaro A. and Civas A. (2009) The role of differential expression of human interferon--a genes in antiviral immunity. *Cytokine Growth Factor Rev* 20, 283-95.
Gilley D., Lee M. S. and Blackburn E. H. (1995) Altering specific telomerase RNA template residues affects active site function. *Genes Dev* 9, 2214-26.

Gilliet M., Boonstra A., Paturel C., Antonenko S., Xu X. L., Trinchieri G., O'Garra A. and Liu Y. J. (2002) The development of murine plasmacytoid dendritic cell precursors is differentially regulated by FLT3-ligand and granulocyte/macrophage colony-stimulating factor. *J Exp Med* 195, 953-8.

Gitlin L., Barchet W., Gilfillan S., Cella M., Beutler B., Flavell R. A., Diamond M. S. and Colonna M. (2006) Essential role of mda-5 in type I IFN responses to polyriboinosinic:polyribocytidylic acid and encephalomyocarditis picornavirus. *Proc Natl Acad Sci U S A* 103, 8459-64.

Gorden K. K., Qiu X. X., Binsfeld C. C., Vasilakos J. P. and Alkan S. S. (2006) Cutting edge: activation of murine TLR8 by a combination of imidazoquinoline immune response modifiers and polyT oligodeoxynucleotides. *J Immunol* 177, 6584-7.

Goutieres F. (2005) Aicardi-Goutieres syndrome. *Brain Dev.* 27, 201-206.

Goutieres F., Aicardi J., Barth P. G. and Lebon P. (1998) Aicardi-Goutieres syndrome: an update and results of interferon-alpha studies. *Ann.Neurol.* 44, 900-907.

Graham R. R., Kozyrev S. V., Baechler E. C., Reddy M. V., Plenge R. M., Bauer J. W., Ortmann W. A., Koeuth T., Gonzalez Escribano M. F., Pons-Estel B., Petri M., Daly M., Gregersen P. K., Martin J., Altshuler D., Behrens T. W. and Alarcon-Riquelme M. E. (2006) A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet* 38, 550-5.

Gray D. M. and Ratliff R. L. (1975) Circular dichroism spectra of poly[d(AC):d(GT)], poly[r(AC):r(GU)], and hybrids poly[d(AC):r(GU)] and poly[r(AC):d(GT)] in the presence of ethanol. *Biopolymers* 14, 487-98.

Gray R. C., Kuchtey J. and Harding C. V. (2007) CpG-B ODNs potently induce low levels of IFN-alphabeta and induce IFN-alphabeta-dependent MHC-I cross-presentation in DCs as effectively as CpG-A and CpG-C ODNs. *J Leukoc Biol* 81, 1075-85.

Guiducci C., Coffman R. L. and Barrat F. J. (2009) Signalling pathways leading to IFN-alpha production in human plasmacytoid dendritic cell and the possible use of agonists or antagonists of TLR7 and TLR9 in clinical indications. *J Intern Med* 265, 43-57.

Guiducci C., Ott G., Chan J. H., Damon E., Calacsan C., Matray T., Lee K. D., Coffman R. L. and Barrat F. J. (2006) Properties regulating the nature of the plasmacytoid dendritic cell response to Toll-like receptor 9 activation. *J Exp Med* 203, 1999-2008.

Gunnarsson G. H., Gudmundsson B., Thormar H. G., Alfredsson A. and Jonsson J. J. (2006a) Two-dimensional strandness-dependent electrophoresis. *Nat Protoc* 1, 3011-8.

Gunnarsson G. H., Gudmundsson B., Thormar H. G., Alfredsson A. and Jonsson J. J. (2006b) Two-dimensional strandness-dependent electrophoresis: a method to characterize single-stranded DNA, double-stranded DNA, and RNA-DNA hybrids in complex samples. *Anal Biochem* 350, 120-7.

Gururajan M., Jacob J. and Pulendran B. (2007) Toll-like receptor expression and responsiveness of distinct murine splenic and mucosal B-cell subsets. *PLoS One* 2, e863.

Gyi J. I., Conn G. L., Lane A. N. and Brown T. (1996) Comparison of the thermodynamic stabilities and solution conformations of DNA.RNA hybrids containing purine-rich and pyrimidine-rich strands with DNA and RNA duplexes. *Biochemistry* 35, 12538-48.

Gyi J. I., Lane A. N., Conn G. L. and Brown T. (1998) The orientation and dynamics of the C2'-OH and hydration of RNA and DNA.RNA hybrids. *Nucleic Acids Res* 26, 3104-10.

Gyllensten U. B. and Erlich H. A. (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc Natl Acad Sci U S A* 85, 7652-6.

Haas T., Metzger J., Schmitz F., Heit A., Muller T., Latz E. and Wagner H. (2008) The DNA sugar backbone 2' deoxyribose determines toll-like receptor 9 activation. *Immunity* 28, 315-23.

Haas T., Schmitz F., Heit A. and Wagner H. (2009) Sequence independent interferon-alpha induction by multimerized phosphodiester DNA depends on spatial regulation of Toll-like receptor-9 activation in plasmacytoid dendritic cells. *Immunology* 126, 290-8.

Hacker H., Mischak H., Miethke T., Liptay S., Schmid R., Sparwasser T., Heeg K., Lipford G. B. and Wagner H. (1998) CpG-DNA-specific activation of antigen-presenting cells requires stress kinase activity and is preceded by non-specific endocytosis and endosomal maturation. *EMBO J* 17, 6230-40.

Hall K. B. and McLaughlin L. W. (1991) Thermodynamic and structural properties of pentamer DNA.DNA, RNA.RNA, and DNA.RNA duplexes of identical sequence. *Biochemistry* 30, 10606-13.
Hardy M. P., Owczarek C. M., Jermiin L. S., Ejdeback M. and Hertzog P. J. (2004) Characterization of the type I interferon locus and identification of novel genes. *Genomics* 84, 331-45.
Harrington J. J. and Lieber M. R. (1994) The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J* 13, 1235-46.
Haruki M., Hayashi K., Kochi T., Muroya A., Koga Y., Morikawa M., Imanaka T. and Kanaya S. (1998) Gene cloning and characterization of recombinant RNase HII from a hyperthermophilic archaeon. *J.Bacteriol.* 180, 6207-6214.
Hausen P. and Stein H. (1970) Ribonuclease H. An enzyme degrading the RNA moiety of DNA-RNA hybrids. *Eur J Biochem* 14, 278-83.
Hausmann S., Marq J. B., Tapparel C., Kolakofsky D. and Garcin D. (2008) RIG-I and dsRNA-induced IFNbeta activation. *PLoS ONE* 3, e3965.
Heil F., Hemmi H., Hochrein H., Ampenberger F., Kirschning C., Akira S., Lipford G., Wagner H. and Bauer S. (2004) Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* 303, 1526-9.
Hemmi H., Kaisho T., Takeuchi O., Sato S., Sanjo H., Hoshino K., Horiuchi T., Tomizawa H., Takeda K. and Akira S. (2002) Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. *Nat Immunol* 3, 196-200.
Hemmi H., Takeuchi O., Kawai T., Kaisho T., Sato S., Sanjo H., Matsumoto M., Hoshino K., Wagner H., Takeda K. and Akira S. (2000) A Toll-like receptor recognizes bacterial DNA. *Nature* 408, 740-5.
Hirayama H., Tamaoka J. and Horikoshi K. (1996) Improved immobilization of DNA to microwell plates for DNA-DNA hybridization. *Nucleic Acids Res* 24, 4098-9.
Ho Y., Gruhler A., Heilbut A., Bader G. D., Moore L., Adams S. L., Millar A., Taylor P., Bennett K., Boutilier K., Yang L., Wolting C., Donaldson I., Schandorff S., Shewnarane J., Vo M., Taggart J., Goudreault M., Muskat B., Alfarano C., Dewar D., Lin Z., Michalickova K., Willems A. R., Sassi H., Nielsen P. A., Rasmussen K. J., Andersen J. R., Johansen L. E., Hansen L. H., Jespersen H., Podtelejnikov A., Nielsen E., Crawford J., Poulsen V., Sorensen B. D., Matthiesen J., Hendrickson R. C., Gleeson F., Pawson T., Moran M. F., Durocher D., Mann M., Hogue C. W., Figeys D. and Tyers M. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180-3.
Hoene V., Peiser M. and Wanner R. (2006) Human monocyte-derived dendritic cells express TLR9 and react directly to the CpG-A oligonucleotide D19. *J Leukoc Biol* 80, 1328-36.
Hoffmann J. A., Kafatos F. C., Janeway C. A. and Ezekowitz R. A. (1999) Phylogenetic perspectives in innate immunity. *Science* 284, 1313-8.
Hogrefe H. H., Hogrefe R. I., Walder R. Y. and Walder J. A. (1990) Kinetic analysis of Escherichia coli RNase H using DNA-RNA-DNA/DNA substrates. *J Biol Chem* 265, 5561-6.
Holm C., Goto T., Wang J. C. and Botstein D. (1985) DNA topoisomerase II is required at the time of mitosis in yeast. *Cell* 41, 553-63.
Honda K., Ohba Y., Yanai H., Negishi H., Mizutani T., Takaoka A., Taya C. and Taniguchi T. (2005) Spatiotemporal regulation of MyD88-IRF-7 signalling for robust type-I interferon induction. *Nature* 434, 1035-40.
Honda K., Takaoka A. and Taniguchi T. (2006) Type I interferon gene induction by the interferon regulatory factor family of transcription factors. *Immunity* 25, 349-60.
Horard B. and Gilson E. (2008) Telomeric RNA enters the game. *Nat Cell Biol* 10, 113-5.
Hornung V., Ablasser A., Charrel-Dennis M., Bauernfeind F., Horvath G., Caffrey D. R., Latz E. and Fitzgerald K. A. (2009) AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* 458, 514-8.
Hornung V., Ellegast J., Kim S., Brzozka K., Jung A., Kato H., Poeck H., Akira S., Conzelmann K. K., Schlee M., Endres S. and Hartmann G. (2006) 5'-Triphosphate RNA is the ligand for RIG-I. *Science* 314, 994-7.
Hornung V., Rothenfusser S., Britsch S., Krug A., Jahrsdorfer B., Giese T., Endres S. and Hartmann G. (2002) Quantitative expression of toll-like receptor 1-10 mRNA in cellular subsets of human peripheral blood mononuclear cells and sensitivity to CpG oligodeoxynucleotides. *J Immunol* 168, 4531-7.

Hornung V., Schlender J., Guenthner-Biller M., Rothenfusser S., Endres S., Conzelmann K. K. and Hartmann G. (2004) Replication-dependent potent IFN-alpha induction in human plasmacytoid dendritic cells by a single-stranded RNA virus. *J Immunol* 173, 5935-43.
Horton N. C. and Finzel B. C. (1996) The structure of an RNA/DNA hybrid: a substrate of the ribonuclease activity of HIV-1 reverse transcriptase. *J Mol Biol* 264, 521-33.
Hoss M., Robins P., Naven T. J., Pappin D. J., Sgouros J. and Lindahl T. (1999) A human DNA editing enzyme homologous to the Escherichia coli DnaQ/MutD protein. *Embo J* 18, 3868-75.
Hsu S. T., Chou M. T. and Cheng J. W. (2000) The solution structure of [d(CGC)r(aaa)d(TTTGCG)](2): hybrid junctions flanked by DNA duplexes. *Nucleic Acids Res* 28, 1322-31.
Hu Z., Zhang A., Storz G., Gottesman S. and Leppla S. H. (2006) An antibody-based microarray assay for small RNA detection. *Nucleic Acids Res* 34, e52.
Huang D., Wujek J., Kidd G., He T. T., Cardona A., Sasse M. E., Stein E. J., Kish J., Tani M., Charo I. F., Proudfoot A. E., Rollins B. J., Handel T. and Ransohoff R. M. (2005) Chronic expression of monocyte chemoattractant protein-1 in the central nervous system causes delayed encephalopathy and impaired microglial function in mice. *FASEB J* 19, 761-72.
Huang F. T., Yu K., Hsieh C. L. and Lieber M. R. (2006) Downstream boundary of chromosomal R-loops at murine switch regions: implications for the mechanism of class switch recombination. *Proc Natl Acad Sci U S A* 103, 5030-5.
Huang Y., Chen C. and Russu I. M. (2009) Dynamics and stability of individual base pairs in two homologous RNA-DNA hybrids. *Biochemistry* 48, 3988-97.
Huertas P. and Aguilera A. (2003) Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell* 12, 711-21.
Hung S. H., Yu Q., Gray D. M. and Ratliff R. L. (1994) Evidence from CD spectra that d(purine).r(pyrimidine) and r(purine).d(pyrimidine) hybrids are in different structural classes. *Nucleic Acids Res* 22, 4326-34.
Ii M. and Brill S. J. (2005) Roles of SGS1, MUS81, and RAD51 in the repair of lagging-strand replication defects in Saccharomyces cerevisiae. *Curr Genet* 48, 213-25.
Imaeda A. B., Watanabe A., Sohail M. A., Mahmood S., Mohamadnejad M., Sutterwala F. S., Flavell R. A. and Mehal W. Z. (2009) Acetaminophen-induced hepatotoxicity in mice is dependent on Tlr9 and the Nalp3 inflammasome. *J Clin Invest* 119, 305-14.
Ioannou Y. and Isenberg D. A. (2000) Current evidence for the induction of autoimmune rheumatic manifestations by cytokine therapy. *Arthritis Rheum* 43, 1431-42.
Ishii K. J., Coban C., Kato H., Takahashi K., Torii Y., Takeshita F., Ludwig H., Sutter G., Suzuki K., Hemmi H., Sato S., Yamamoto M., Uematsu S., Kawai T., Takeuchi O. and Akira S. (2006) A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nat Immunol* 7, 40-8.
Ishii K. J., Kawagoe T., Koyama S., Matsui K., Kumar H., Kawai T., Uematsu S., Takeuchi O., Takeshita F., Coban C. and Akira S. (2008) TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. *Nature* 451, 725-729.
Ishikawa H. and Barber G. N. (2008) STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* 455, 674-8.
Ishikawa H., Ma Z. and Barber G. N. (2009) STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* 461, 788-92.
Itaya M. (1990) Isolation and characterization of a second RNase H (RNase HII) of Escherichia coli K-12 encoded by the rnhB gene. *Proc Natl Acad Sci U S A* 87, 8587-91.
Itoh T. and Tomizawa J. (1980) Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proc Natl Acad Sci U S A* 77, 2450-4.
Iwasaki A. and Medzhitov R. (2010) Regulation of adaptive immunity by the innate immune system. *Science* 327, 291-5.
Izaguirre A., Barnes B. J., Amrute S., Yeow W. S., Megjugorac N., Dai J., Feng D., Chung E., Pitha P. M. and Fitzgerald-Bocarsly P. (2003) Comparative analysis of IRF and IFN-alpha expression in human plasmacytoid and monocyte-derived dendritic cells. *J Leukoc Biol* 74, 1125-38.
Jaks E., Gavutis M., Uze G., Martal J. and Piehler J. (2007) Differential receptor subunit affinities of type I interferons govern differential signal activation. *J Mol Biol* 366, 525-39.

Jeong H. S., Backlund P. S., Chen H. C., Karavanov A. A. and Crouch R. J. (2004) RNase H2 of Saccharomyces cerevisiae is a complex of three proteins. *Nucleic Acids Res.* 32, 407-414.
Joyce C. M. (1997) Choosing the right sugar: how polymerases select a nucleotide substrate. *Proc.Natl.Acad.Sci.U.S.A* 94, 1619-1622.
Jurk M., Heil F., Vollmer J., Schetter C., Krieg A. M., Wagner H., Lipford G. and Bauer S. (2002) Human TLR7 or TLR8 independently confer responsiveness to the antiviral compound R-848. *Nat Immunol* 3, 499.
Kadowaki N., Ho S., Antonenko S., Malefyt R. W., Kastelein R. A., Bazan F. and Liu Y. J. (2001) Subsets of human dendritic cell precursors express different toll-like receptors and respond to different microbial antigens. *J Exp Med* 194, 863-9.
Kaiser W. J., Upton J. W. and Mocarski E. S. (2008) Receptor-interacting protein homotypic interaction motif-dependent control of NF-kappa B activation via the DNA-dependent activator of IFN regulatory factors. *J Immunol* 181, 6427-34.
Kaltenboeck B., Spatafora J. W., Zhang X., Kousoulas K. G., Blackwell M. and Storz J. (1992) Efficient production of single-stranded DNA as long as 2 kb for sequencing of PCR-amplified DNA. *Biotechniques* 12, 164, 166, 168-71.
Kang D. C., Gopalkrishnan R. V., Wu Q., Jankowsky E., Pyle A. M. and Fisher P. B. (2002) mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. *Proc Natl Acad Sci U S A* 99, 637-42.
Kao H. I. and Bambara R. A. (2003) The protein components and mechanism of eukaryotic Okazaki fragment maturation. *Crit Rev Biochem Mol Biol* 38, 433-52.
Karayel E., Burckstummer T., Bilban M., Durnberger G., Weitzer S., Martinez J. and Superti-Furga G. (2009) The TLR-independent DNA recognition pathway in murine macrophages: Ligand features and molecular signature. *Eur J Immunol* 39, 1929-36.
Kataoka I., Shinagawa K., Shiro Y., Okamoto S., Watanabe R., Mori T., Ito D. and Harada M. (2002) Multiple sclerosis associated with interferon-alpha therapy for chronic myelogenous leukemia. *Am J Hematol* 70, 149-53.
Katayanagi K., Miyagawa M., Matsushima M., Ishikawa M., Kanaya S., Ikehara M., Matsuzaki T. and Morikawa K. (1990) Three-dimensional structure of ribonuclease H from E. coli. *Nature* 347, 306-9.
Kato H., Sato S., Yoneyama M., Yamamoto M., Uematsu S., Matsui K., Tsujimura T., Takeda K., Fujita T., Takeuchi O. and Akira S. (2005) Cell type-specific involvement of RIG-I in antiviral response. *Immunity* 23, 19-28.
Kato H., Takeuchi O., Mikamo-Satoh E., Hirai R., Kawai T., Matsushita K., Hiiragi A., Dermody T. S., Fujita T. and Akira S. (2008) Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *J Exp Med* 205, 1601-10.
Kato H., Takeuchi O., Sato S., Yoneyama M., Yamamoto M., Matsui K., Uematsu S., Jung A., Kawai T., Ishii K. J., Yamaguchi O., Otsu K., Tsujimura T., Koh C. S., Reis e Sousa C., Matsuura Y., Fujita T. and Akira S. (2006) Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 441, 101-5.
Kawai T. and Akira S. (2006) Innate immune recognition of viral infection. *Nat.Immunol.* 7, 131-137.
Kawai T. and Akira S. (2009) The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int Immunol* 21, 317-37.
Kawai T., Takahashi K., Sato S., Coban C., Kumar H., Kato H., Ishii K. J., Takeuchi O. and Akira S. (2005) IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat Immunol* 6, 981-8.
Kawane K., Fukuyama H., Kondoh G., Takeda J., Ohsawa Y., Uchiyama Y. and Nagata S. (2001) Requirement of DNase II for definitive erythropoiesis in the mouse fetal liver. *Science* 292, 1546-9.
Kawane K., Fukuyama H., Yoshida H., Nagase H., Ohsawa Y., Uchiyama Y., Okada K., Iida T. and Nagata S. (2003) Impaired thymic development in mouse embryos deficient in apoptotic DNA degradation. *Nat Immunol* 4, 138-44.
Kawane K., Ohtani M., Miwa K., Kizawa T., Kanbara Y., Yoshioka Y., Yoshikawa H. and Nagata S. (2006) Chronic polyarthritis caused by mammalian DNA that escapes from degradation in macrophages. *Nature* 443, 998-1002.

Keller W. and Crouch R. (1972) Degradation of DNA RNA hybrids by ribonuclease H and DNA polymerases of cellular and viral origin. *Proc Natl Acad Sci U S A* 69, 3360-4.
Kerkmann M., Rothenfusser S., Hornung V., Towarowski A., Wagner M., Sarris A., Giese T., Endres S. and Hartmann G. (2003) Activation with CpG-A and CpG-B oligonucleotides reveals two distinct regulatory pathways of type I IFN synthesis in human plasmacytoid dendritic cells. *J Immunol* 170, 4465-74.
Kikuchi K., Taniguchi Y., Hatanaka A., Sonoda E., Hochegger H., Adachi N., Matsuzaki Y., Koyama H., van Gent D. C., Jasin M. and Takeda S. (2005) Fen-1 facilitates homologous recombination by removing divergent sequences at DNA break ends. *Mol Cell Biol* 25, 6948-55.
Kim I., McKenna S. A., Viani Puglisi E. and Puglisi J. D. (2007) Rapid purification of RNAs using fast performance liquid chromatography (FPLC). *RNA* 13, 289-94.
Kim N. W., Piatyszek M. A., Prowse K. R., Harley C. B., West M. D., Ho P. L., Coviello G. M., Wright W. E., Weinrich S. L. and Shay J. W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011-5.
Kim S., Kim N., Dong B., Boren D., Lee S. A., Das Gupta J., Gaughan C., Klein E. A., Lee C., Silverman R. H. and Chow S. A. (2008a) Integration site preference of xenotropic murine leukemia virus-related virus, a new human retrovirus associated with prostate cancer. *J Virol* 82, 9964-77.
Kim T., Pazhoor S., Bao M., Zhang Z., Hanabuchi S., Facchinetti V., Bover L., Plumas J., Chaperot L., Qin J. and Liu Y. J. (2010) Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells. *Proc Natl Acad Sci U S A* 107, 15181-6.
Kim Y. M., Brinkmann M. M., Paquet M. E. and Ploegh H. L. (2008b) UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* 452, 234-8.
Kinney J. S., Viscidi R. P., Vonderfecht S. L., Eiden J. J. and Yolken R. H. (1989) Monoclonal antibody assay for detection of double-stranded RNA and application for detection of group A and non-group A rotaviruses. *J Clin Microbiol* 27, 6-12.
Kireeva M. L., Komissarova N. and Kashlev M. (2000a) Overextended RNA:DNA hybrid as a negative regulator of RNA polymerase II processivity. *J Mol Biol* 299, 325-35.
Kireeva M. L., Komissarova N., Waugh D. S. and Kashlev M. (2000b) The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J Biol Chem* 275, 6530-6.
Kitagawa Y. and Stollar B. D. (1982) Comparison of poly(A).poly(dT) and poly(I).poly(dC) as immunogens for the induction of antibodies to RNA-DNA hybrids. *Mol Immunol* 19, 413-20.
Klarmann G. J., Schauber C. A. and Preston B. D. (1993) Template-directed pausing of DNA synthesis by HIV-1 reverse transcriptase during polymerization of HIV-1 sequences in vitro. *J Biol Chem* 268, 9793-802.
Klein R. S. (2004) Regulation of neuroinflammation: the role of CXCL10 in lymphocyte infiltration during autoimmune encephalomyelitis. *J Cell Biochem* 92, 213-22.
Kleinman M. E., Yamada K., Takeda A., Chandrasekaran V., Nozaki M., Baffi J. Z., Albuquerque R. J., Yamasaki S., Itaya M., Pan Y., Appukuttan B., Gibbs D., Yang Z., Kariko K., Ambati B. K., Wilgus T. A., DiPietro L. A., Sakurai E., Zhang K., Smith J. R., Taylor E. W. and Ambati J. (2008) Sequence- and target-independent angiogenesis suppression by siRNA via TLR3. *Nature* 452, 591-7.
Klune J. R., Dhupar R., Cardinal J., Billiar T. R. and Tsung A. (2008) HMGB1: endogenous danger signaling. *Mol Med* 14, 476-84.
Komatsuda A., Wakui H., Iwamoto K., Ozawa M., Togashi M., Masai R., Maki N., Hatakeyama T. and Sawada K. (2008) Up-regulated expression of Toll-like receptors mRNAs in peripheral blood mononuclear cells from patients with systemic lupus erythematosus. *Clin Exp Immunol* 152, 482-7.
Komuro A. and Horvath C. M. (2006) RNA- and virus-independent inhibition of antiviral signaling by RNA helicase LGP2. *J Virol* 80, 12332-42.
Kotenko S. V., Gallagher G., Baurin V. V., Lewis-Antes A., Shen M., Shah N. K., Langer J. A., Sheikh F., Dickensheets H. and Donnelly R. P. (2003) IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat Immunol* 4, 69-77.

Kozyrev S. V., Lewen S., Reddy P. M., Pons-Estel B., Witte T., Junker P., Laustrup H., Gutierrez C., Suarez A., Francisca Gonzalez-Escribano M., Martin J. and Alarcon-Riquelme M. E. (2007) Structural insertion/deletion variation in IRF5 is associated with a risk haplotype and defines the precise IRF5 isoforms expressed in systemic lupus erythematosus. *Arthritis Rheum* 56, 1234-41.

Krangel M. S. (2009) Mechanics of T cell receptor gene rearrangement. *Curr Opin Immunol* 21, 133-9.

Krieg A. M. (2002) CpG motifs in bacterial DNA and their immune effects. *Annu Rev Immunol* 20, 709-60.

Krieser R. J., MacLea K. S., Longnecker D. S., Fields J. L., Fiering S. and Eastman A. (2002) Deoxyribonuclease IIalpha is required during the phagocytic phase of apoptosis and its loss causes perinatal lethality. *Cell Death Differ* 9, 956-62.

Krug A., Luker G. D., Barchet W., Leib D. A., Akira S. and Colonna M. (2004) Herpes simplex virus type 1 activates murine natural interferon-producing cells through toll-like receptor 9. *Blood* 103, 1433-7.

Krug A., Rothenfusser S., Hornung V., Jahrsdorfer B., Blackwell S., Ballas Z. K., Endres S., Krieg A. M. and Hartmann G. (2001) Identification of CpG oligonucleotide sequences with high induction of IFN-alpha/beta in plasmacytoid dendritic cells. *Eur J Immunol* 31, 2154-63.

Kujau M. J. and Wolfl S. (1997) Efficient preparation of single-stranded DNA for in vitro selection. *Mol Biotechnol* 7, 333-5.

Kumar H., Kawai T., Kato H., Sato S., Takahashi K., Coban C., Yamamoto M., Uematsu S., Ishii K. J., Takeuchi O. and Akira S. (2006) Essential role of IPS-1 in innate immune responses against RNA viruses. *J Exp Med* 203, 1795-803.

Kurt-Jones E. A., Sandor F., Ortiz Y., Bowen G. N., Counter S. L., Wang T. C. and Finberg R. W. (2004) Use of murine embryonic fibroblasts to define Toll-like receptor activation and specificity. *J Endotoxin Res* 10, 419-24.

Kyogoku C., Langefeld C. D., Ortmann W. A., Lee A., Selby S., Carlton V. E., Chang M., Ramos P., Baechler E. C., Batliwalla F. M., Novitzke J., Williams A. H., Gillett C., Rodine P., Graham R. R., Ardlie K. G., Gaffney P. M., Moser K. L., Petri M., Begovich A. B., Gregersen P. K. and Behrens T. W. (2004) Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am J Hum Genet* 75, 504-7.

Lacotte S., Brun S., Muller S. and Dumortier H. (2009) CXCR3, inflammation, and autoimmune diseases. *Ann N Y Acad Sci* 1173, 310-7.

Ladoyanni E. and Nambi R. (2005) Psoriasis exacerbated by interferon-alpha in a patient with chronic myeloid leukemia. *J Drugs Dermatol* 4, 221-2.

Lai L., Yokota H., Hung L. W., Kim R. and Kim S. H. (2000) Crystal structure of archaeal RNase HII: a homologue of human major RNase H. *Structure* 8, 897-904.

Lande R. and Gilliet M. (2010) Plasmacytoid dendritic cells: key players in the initiation and regulation of immune responses. *Ann N Y Acad Sci* 1183, 89-103.

Lande R., Gregorio J., Facchinetti V., Chatterjee B., Wang Y. H., Homey B., Cao W., Wang Y. H., Su B., Nestle F. O., Zal T., Mellman I., Schroder J. M., Liu Y. J. and Gilliet M. (2007) Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide. *Nature* 449, 564-9.

Lanzi G., Fazzi E. and D'Arrigo S. (2002) Aicardi-Goutieres syndrome: a description of 21 new cases and a comparison with the literature. *Eur.J.Paediatr.Neurol.* 6 Suppl A, A9-22.

Lanzi G., Fazzi E., D'Arrigo S., Orcesi S., Maraucci I., Uggetti C., Bertini E. and Lebon P. (2005) The natural history of Aicardi-Goutieres syndrome: follow-up of 11 Italian patients. *Neurology* 64, 1621-4.

Latz E., Schoenemeyer A., Visintin A., Fitzgerald K. A., Monks B. G., Knetter C. F., Lien E., Nilsen N. J., Espevik T. and Golenbock D. T. (2004) TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nat Immunol* 5, 190-8.

Lau C. M., Broughton C., Tabor A. S., Akira S., Flavell R. A., Mamula M. J., Christensen S. R., Shlomchik M. J., Viglianti G. A., Rifkin I. R. and Marshak-Rothstein A. (2005) RNA-associated autoantigens activate B cells by combined B cell antigen receptor/Toll-like receptor 7 engagement. *J Exp Med* 202, 1171-7.

Lauretti F., Lucas de Melo F., Benati F. J., de Mello Volotao E., Santos N., Linhares R. E. and Nozawa C. (2003) Use of acridine orange staining for the detection of rotavirus RNA in polyacrylamide gels. *J Virol Methods* 114, 29-35.

Le Bon A., Schiavoni G., D'Agostino G., Gresser I., Belardelli F. and Tough D. F. (2001) Type i interferons potently enhance humoral immunity and can promote isotype switching by stimulating dendritic cells in vivo. *Immunity* 14, 461-70.
Leadbetter E. A., Rifkin I. R., Hohlbaum A. M., Beaudette B. C., Shlomchik M. J. and Marshak-Rothstein A. (2002) Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. *Nature* 416, 603-7.
Lebon P., Badoual J., Ponsot G., Goutieres F., Hemeury-Cukier F. and Aicardi J. (1988) Intrathecal synthesis of interferon-alpha in infants with progressive familial encephalopathy. *J.Neurol.Sci.* 84, 201-208.
Lebon P., Meritet J. F., Krivine A. and Rozenberg F. (2002) Interferon and Aicardi-Goutieres syndrome. *Eur J Paediatr Neurol* 6 Suppl A, A47-53.
Lee-Kirsch M. A., Chowdhury D., Harvey S., Gong M., Senenko L., Engel K., Pfeiffer C., Hollis T., Gahr M., Perrino F. W., Lieberman J. and Hubner N. (2007a) A mutation in TREX1 that impairs susceptibility to granzyme A-mediated cell death underlies familial chilblain lupus. *J Mol Med* 85, 531-7.
Lee-Kirsch M. A., Gong M., Chowdhury D., Senenko L., Engel K., Lee Y. A., de Silva U., Bailey S. L., Witte T., Vyse T. J., Kere J., Pfeiffer C., Harvey S., Wong A., Koskenmies S., Hummel O., Rohde K., Schmidt R. E., Dominiczak A. F., Gahr M., Hollis T., Perrino F. W., Lieberman J. and Hubner N. (2007b) Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. *Nat Genet* 39, 1065-7.
Lee-Kirsch M. A., Gong M., Schulz H., Ruschendorf F., Stein A., Pfeiffer C., Ballarini A., Gahr M., Hubner N. and Linne M. (2006) Familial chilblain lupus, a monogenic form of cutaneous lupus erythematosus, maps to chromosome 3p. *Am J Hum Genet* 79, 731-7.
Lee D. Y. and Clayton D. A. (1996) Properties of a primer RNA-DNA hybrid at the mouse mitochondrial DNA leading-strand origin of replication. *J Biol Chem* 271, 24262-9.
Lee H. K., Lund J. M., Ramanathan B., Mizushima N. and Iwasaki A. (2007) Autophagy-dependent viral recognition by plasmacytoid dendritic cells. *Science* 315, 1398-401.
Lee S. H., Pan Z. Q., Kwong A. D., Burgers P. M. and Hurwitz J. (1991) Synthesis of DNA by DNA polymerase epsilon in vitro. *J Biol Chem* 266, 22707-17.
Lee Y. N. and Bieniasz P. D. (2007) Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog* 3, e10.
Lehtinen D. A., Harvey S., Mulcahy M. J., Hollis T. and Perrino F. W. (2008) The TREX1 double-stranded DNA degradation activity is defective in dominant mutations associated with autoimmune disease. *J Biol Chem* 283, 31649-56.
Leifer C. A., Kennedy M. N., Mazzoni A., Lee C., Kruhlak M. J. and Segal D. M. (2004) TLR9 is localized in the endoplasmic reticulum prior to stimulation. *J Immunol* 173, 1179-83.
Lesnik E. A. and Freier S. M. (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry* 34, 10807-15.
Li N., Zhang W. and Cao X. (2000) Identification of human homologue of mouse IFN-gamma induced protein from human dendritic cells. *Immunol Lett* 74, 221-4.
Li W. M., Barnes T. and Lee C. H. (2010) Endoribonucleases--enzymes gaining spotlight in mRNA metabolism. *FEBS J* 277, 627-41.
Li X. and Manley J. L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* 122, 365-78.
Li X. and Manley J. L. (2006) Cotranscriptional processes and their influence on genome stability. *Genes Dev* 20, 1838-47.
Li X. L., Boyanapalli M., Weihua X., Kalvakolanu D. V. and Hassel B. A. (1998) Induction of interferon synthesis and activation of interferon-stimulated genes by liposomal transfection reagents. *J Interferon Cytokine Res* 18, 947-52.
Li X. L., Ezelle H. J., Kang T. J., Zhang L., Shirey K. A., Harro J., Hasday J. D., Mohapatra S. K., Crasta O. R., Vogel S. N., Cross A. S. and Hassel B. A. (2008) An essential role for the antiviral endoribonuclease, RNase-L, in antibacterial immunity. *Proc Natl Acad Sci U S A* 105, 20816-21.
Liao W., Bao Z., Cheng C., Mok Y. K. and Wong W. S. (2008) Dendritic cell-derived interferon-gamma-induced protein mediates tumor necrosis factor-alpha stimulation of human lung fibroblasts. *Proteomics* 8, 2640-50.

Licht R., Dieker J. W., Jacobs C. W., Tax W. J. and Berden J. H. (2004) Decreased phagocytosis of apoptotic cells in diseased SLE mice. *J Autoimmun* 22, 139-45.
Lin Y., Dent S. Y., Wilson J. H., Wells R. D. and Napierala M. (2010) R loops stimulate genetic instability of CTG.CAG repeats. *Proc Natl Acad Sci U S A* 107, 692-7.
Lindahl T., Barnes D. E., Yang Y. G. and Robins P. (2009) Biochemical properties of mammalian TREX1 and its association with DNA replication and inherited inflammatory disease. *Biochem Soc Trans* 37, 535-8.
Lindahl T., Gally J. A. and Edelman G. M. (1969) Properties of deoxyribonuclease 3 from mammalian tissues. *J Biol Chem* 244, 5014-9.
Lippmann J., Rothenburg S., Deigendesch N., Eitel J., Meixenberger K., van Laak V., Slevogt H., N'Guessan P D., Hippenstiel S., Chakraborty T., Flieger A., Suttorp N. and Opitz B. (2008) IFNbeta responses induced by intracellular bacteria or cytosolic DNA in different human cells do not require ZBP1 (DLM-1/DAI). *Cell Microbiol* 10, 2579-88.
Liu W., Liang S. L., Liu H., Silverman R. and Zhou A. (2007) Tumour suppressor function of RNase L in a mouse model. *Eur J Cancer* 43, 202-9.
Livak K. J. and Schmittgen T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-8.
Loetscher M., Loetscher P., Brass N., Meese E. and Moser B. (1998) Lymphocyte-specific chemokine receptor CXCR3: regulation, chemokine binding and gene localization. *Eur J Immunol* 28, 3696-705.
Lovgren T., Eloranta M. L., Bave U., Alm G. V. and Ronnblom L. (2004) Induction of interferon-alpha production in plasmacytoid dendritic cells by immune complexes containing nucleic acid released by necrotic or late apoptotic cells and lupus IgG. *Arthritis Rheum* 50, 1861-72.
Lukavsky P. J. and Puglisi J. D. (2004) Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides. *RNA* 10, 889-93.
Luke B., Panza A., Redon S., Iglesias N., Li Z. and Lingner J. (2008) The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. *Mol Cell* 32, 465-77.
Lund J., Sato A., Akira S., Medzhitov R. and Iwasaki A. (2003) Toll-like receptor 9-mediated recognition of Herpes simplex virus-2 by plasmacytoid dendritic cells. *J Exp Med* 198, 513-20.
Lund J. M., Alexopoulou L., Sato A., Karow M., Adams N. C., Gale N. W., Iwasaki A. and Flavell R. A. (2004) Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc Natl Acad Sci U S A* 101, 5598-603.
Luster A. D., Unkeless J. C. and Ravetch J. V. (1985) Gamma-interferon transcriptionally regulates an early-response gene containing homology to platelet proteins. *Nature* 315, 672-6.
Lutz M. B., Kukutsch N., Ogilvie A. L., Rossner S., Koch F., Romani N. and Schuler G. (1999) An advanced culture method for generating large quantities of highly pure dendritic cells from mouse bone marrow. *J Immunol Methods* 223, 77-92.
Lyon C. J., Miranda G. A., Piao J. S. and Aguilera R. J. (1996) Characterization of an endonuclease activity which preferentially cleaves the G-rich immunoglobulin switch repeat sequences. *Mol Immunol* 33, 157-69.
Ma W. P. and Crouch R. J. (1996) Escherichia coli RNase HI inhibits murine leukaemia virus reverse transcription in vitro and yeast retrotransposon Ty1 transposition in vivo. *Genes Cells* 1, 581-93.
Macanovic M. and Lachmann P. J. (1997) Measurement of deoxyribonuclease I (DNase) in the serum and urine of systemic lupus erythematosus (SLE)-prone NZB/NZW mice by a new radial enzyme diffusion assay. *Clin Exp Immunol* 108, 220-6.
Madsen B. E., Ramos E. M., Boulard M., Duda K., Overgaard J., Nordsmark M., Wiuf C. and Hansen L. L. (2008) Germline mutation in RNASEL predicts increased risk of head and neck, uterine cervix and breast cancer. *PLoS One* 3, e2492.
Maga G., Villani G., Tillement V., Stucki M., Locatelli G. A., Frouin I., Spadari S. and Hubscher U. (2001) Okazaki fragment processing: modulation of the strand displacement activity of DNA polymerase delta by the concerted action of replication protein A, proliferating cell nuclear antigen, and flap endonuclease-1. *Proc Natl Acad Sci U S A* 98, 14298-303.
Malathi K., Dong B., Gale M., Jr. and Silverman R. H. (2007) Small self-RNA generated by RNase L amplifies antiviral innate immunity. *Nature* 448, 816-9.

Malik H. S. (2005) Ribonuclease H evolution in retrotransposable elements. *Cytogenet Genome Res* 110, 392-401.
Manderson A. P., Botto M. and Walport M. J. (2004) The role of complement in the development of systemic lupus erythematosus. *Annu Rev Immunol* 22, 431-56.
Maraskovsky E., Brasel K., Teepe M., Roux E. R., Lyman S. D., Shortman K. and McKenna H. J. (1996) Dramatic increase in the numbers of functionally mature dendritic cells in Flt3 ligand-treated mice: multiple dendritic cell subpopulations identified. *J Exp Med* 184, 1953-62.
Marie I., Durbin J. E. and Levy D. E. (1998) Differential viral induction of distinct interferon-alpha genes by positive feedback through interferon regulatory factor-7. *Embo J* 17, 6660-9.
Martin D. A. and Elkon K. B. (2005) Autoantibodies make a U-turn: the toll hypothesis for autoantibody specificity. *J Exp Med* 202, 1465-9.
Martinez J., Huang X. and Yang Y. (2010) Toll-like receptor 8-mediated activation of murine plasmacytoid dendritic cells by vaccinia viral DNA. *Proc Natl Acad Sci U S A* 107, 6442-7.
Masukata H. and Tomizawa J. (1990) A mechanism of formation of a persistent hybrid between elongating RNA and template DNA. *Cell* 62, 331-8.
Matsumoto M., Funami K., Tanabe M., Oshiumi H., Shingai M., Seto Y., Yamamoto A. and Seya T. (2003) Subcellular localization of Toll-like receptor 3 in human dendritic cells. *J Immunol* 171, 3154-62.
Matsumoto M., Kikkawa S., Kohase M., Miyake K. and Seya T. (2002) Establishment of a monoclonal antibody against human Toll-like receptor 3 that blocks double-stranded RNA-mediated signaling. *Biochem Biophys Res Commun* 293, 1364-9.
Matzinger P. (1994) Tolerance, danger, and the extended family. *Annu Rev Immunol* 12, 991-1045.
Maxwell I. H., Van Ness J. and Hahn W. E. (1978) Assay of DNA-RNA hybrids by S1 nuclease digestion and adsorption to DEAE-cellulose filters. *Nucleic Acids Res* 5, 2033-8.
Mazur D. J. and Perrino F. W. (1999) Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3'-->5' exonucleases. *J Biol Chem* 274, 19655-60.
Mazur D. J. and Perrino F. W. (2001a) Excision of 3' termini by the Trex1 and TREX2 3'-->5' exonucleases. Characterization of the recombinant proteins. *J Biol Chem* 276, 17022-9.
Mazur D. J. and Perrino F. W. (2001b) Structure and expression of the TREX1 and TREX2 3' --> 5' exonuclease genes. *J Biol Chem* 276, 14718-27.
McGettrick A. F. and O'Neill L. A. (2010) Localisation and trafficking of Toll-like receptors: an important mode of regulation. *Curr Opin Immunol* 22, 20-7.
McMaster G. K. and Carmichael G. G. (1977) Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc Natl Acad Sci U S A* 74, 4835-8.
Means T. K., Latz E., Hayashi F., Murali M. R., Golenbock D. T. and Luster A. D. (2005) Human lupus autoantibody-DNA complexes activate DCs through cooperation of CD32 and TLR9. *J Clin Invest* 115, 407-17.
Melchjorsen J., Jensen S. B., Malmgaard L., Rasmussen S. B., Weber F., Bowie A. G., Matikainen S. and Paludan S. R. (2005) Activation of innate defense against a paramyxovirus is mediated by RIG-I and TLR7 and TLR8 in a cell-type-specific manner. *J Virol* 79, 12944-51.
Menson E. and Lyall H. (2005) Clinical presentation of congenital viral infections *Current Paediatrics* 15, 163-170.
Meurs E., Chong K., Galabru J., Thomas N. S., Kerr I. M., Williams B. R. and Hovanessian A. G. (1990) Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. *Cell* 62, 379-90.
Meyer O. (2009) Interferons and autoimmune disorders. *Joint Bone Spine* 76, 464-73.
Meylan E., Curran J., Hofmann K., Moradpour D., Binder M., Bartenschlager R. and Tschopp J. (2005) Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* 437, 1167-72.
Milhaud P. G., Compagnon B., Bienvenue A. and Philippot J. R. (1992) Interferon production of L929 and HeLa cells enhanced by polyriboinosinic acid-polyribocytidylic acid pH-sensitive liposomes. *Bioconjug Chem* 3, 402-7.
Miller H. I., Riggs A. D. and Gill G. N. (1973) Ribonuclease H (hybrid) in Escherichia coli. Identification and characterization. *J Biol Chem* 248, 2621-4.
Milman G., Langridge R. and Chamberlin M. J. (1967) The structure of a DNA-RNA hybrid. *Proc Natl Acad Sci U S A* 57, 1804-10.

Misquitta L., Wei Q. and Paterson B. M. (2008) Preparation of Double-Stranded RNA for *Drosophila* RNA Interference (RNAi). *Cold Spring Harb. Protoc.* doi:10.1101/pdb.prot4916.
Mitchell M., Gillis A., Futahashi M., Fujiwara H. and Skordalakes E. (2010) Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat Struct Mol Biol* 17, 513-8.
Mogensen K. E., Lewerenz M., Reboul J., Lutfalla G. and Uze G. (1999) The type I interferon receptor: structure, function, and evolution of a family business. *J Interferon Cytokine Res* 19, 1069-98.
Molling K., Bolognesi D. P., Bauer H., Busen W., Plassmann H. W. and Hausen P. (1971) Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat New Biol* 234, 240-3.
Morita M., Stamp G., Robins P., Dulic A., Rosewell I., Hrivnak G., Daly G., Lindahl T. and Barnes D. E. (2004) Gene-targeted mice lacking the Trex1 (DNase III) 3'-->5' DNA exonuclease develop inflammatory myocarditis. *Mol Cell Biol* 24, 6719-27.
Munoz L. E., Janko C., Grossmayer G. E., Frey B., Voll R. E., Kern P., Kalden J. R., Schett G., Fietkau R., Herrmann M. and Gaipl U. S. (2009) Remnants of secondarily necrotic cells fuel inflammation in systemic lupus erythematosus. *Arthritis Rheum* 60, 1733-1742.
Murali A., Li X., Ranjith-Kumar C. T., Bhardwaj K., Holzenburg A., Li P. and Kao C. C. (2008) Structure and function of LGP2, a DEX(D/H) helicase that regulates the innate immunity response. *J Biol Chem* 283, 15825-33.
Muruve D. A., Petrilli V., Zaiss A. K., White L. R., Clark S. A., Ross P. J., Parks R. J. and Tschopp J. (2008) The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response. *Nature* 452, 103-107.
Myong S., Cui S., Cornish P. V., Kirchhofer A., Gack M. U., Jung J. U., Hopfner K. P. and Ha T. (2009) Cytosolic Viral Sensor RIG-I Is a 5'-Triphosphate-Dependent Translocase on Double-Stranded RNA. *Science* 323, 1070-4.
Nadano D., Yasuda T. and Kishi K. (1993) Measurement of deoxyribonuclease I activity in human tissues and body fluids by a single radial enzyme-diffusion method. *Clin Chem* 39, 448-52.
Nagata S., Hanayama R. and Kawane K. (2010) Autoimmunity and the Clearance of Dead Cells. *Cell* 140, 619 - 630.
Naik S. H., Proietto A. I., Wilson N. S., Dakic A., Schnorrer P., Fuchsberger M., Lahoud M. H., O'Keeffe M., Shao Q. X., Chen W. F., Villadangos J. A., Shortman K. and Wu L. (2005) Cutting edge: generation of splenic CD8+ and CD8- dendritic cell equivalents in Fms-like tyrosine kinase 3 ligand bone marrow cultures. *J Immunol* 174, 6592-7.
Nakano S., Morimoto S., Suzuki S., Watanabe T., Amano H. and Takasaki Y. (2010) Up-regulation of the endoplasmic reticulum transmembrane protein UNC93B in the B cells of patients with active systemic lupus erythematosus. *Rheumatology (Oxford)* 49, 876-81.
Nakazato H. (1980) Fractionation and characterization of antibodies elicited phi X174 deoxyribonucleic acid-ribonucleic acid hybrid. *Biochemistry* 19, 2835-40.
Nallagatla S. R., Hwang J., Toroney R., Zheng X., Cameron C. E. and Bevilacqua P. C. (2007) 5'-triphosphate-dependent activation of PKR by RNAs with short stem-loops. *Science* 318, 1455-8.
Napirei M., Karsunky H., Zevnik B., Stephan H., Mannherz H. G. and Moroy T. (2000) Features of systemic lupus erythematosus in Dnase1-deficient mice. *Nat Genet* 25, 177-81.
Nestle F. O., Conrad C., Tun-Kyi A., Homey B., Gombert M., Boyman O., Burg G., Liu Y. J. and Gilliet M. (2005) Plasmacytoid predendritic cells initiate psoriasis through interferon-alpha production. *J Exp Med* 202, 135-43.
Nguyen K. B., Cousens L. P., Doughty L. A., Pien G. C., Durbin J. E. and Biron C. A. (2000) Interferon alpha/beta-mediated inhibition and promotion of interferon gamma: STAT1 resolves a paradox. *Nat.Immunol.* 1, 70-76.
Nichols N. M. and Yue D. (2008) Ribonucleases. *Curr Protoc Mol Biol* Chapter 3, Unit 3 13.
Nicholson A. W. (1999) Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol Rev* 23, 371-90.
Nick McElhinny S. A., Watts B. E., Kumar D., Watt D. L., Lundstrom E. B., Burgers P. M., Johansson E., Chabes A. and Kunkel T. A. (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci U S A* 107, 4949-54.
Nishimura M. and Naito S. (2005) Tissue-specific mRNA expression profiles of human toll-like receptors and related genes. *Biol Pharm Bull* 28, 886-92.

Nishiya T. and DeFranco A. L. (2004) Ligand-regulated chimeric receptor approach reveals] lifespan of murine fibroblasts. *Nat Cell Biol* 5, 741-7.
Pavlov Y. I., Frahm C., Nick McElhinny S. A., Niimi A., Suzuki M. and Kunkel T. A. (2006) Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Curr Biol* 16, 202-7.
Peaston A. E., Evsikov A. V., Graber J. H., de Vries W. N., Holbrook A. E., Solter D. and Knowles B. B. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 7, 597-606.
Perrino F. W., Harvey S., Shaban N. M. and Hollis T. (2009) RNaseH2 mutants that cause Aicardi-Goutieres syndrome are active nucleases. *J Mol Med* 87, 25-30.
Phan A. T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J* 277, 1107-17.
Pichlmair A., Schulz O., Tan C. P., Naslund T. I., Liljestrom P., Weber F. and Reis e Sousa C. (2006) RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* 314, 997-1001.
Pichlmair A., Schulz O., Tan C. P., Rehwinkel J., Kato H., Takeuchi O., Akira S., Way M., Schiavo G. and Reis e Sousa C. (2009) Activation of MDA5 requires higher-order RNA structures generated during virus infection. *J Virol* 83, 10761-9.
Pileur F., Toulme J. J. and Cazenave C. (2000) Eukaryotic ribonucleases HI and HII generate characteristic hydrolytic patterns on DNA-RNA hybrids: further evidence that mitochondrial RNase H is an RNase HII. *Nucleic Acids Res* 28, 3674-83.
Pippig D. A., Hellmuth J. C., Cui S., Kirchhofer A., Lammens K., Lammens A., Schmidt A., Rothenfusser S. and Hopfner K. P. (2009) The regulatory domain of the RIG-I family ATPase LGP2 senses double-stranded RNA. *Nucleic Acids Res* 37, 2014-25.
Poeck H., Bscheider M., Gross O., Finger K., Roth S., Rebsamen M., Hannesschlager N., Schlee M., Rothenfusser S., Barchet W., Kato H., Akira S., Inoue S., Endres S., Peschel C., Hartmann G., Hornung V. and Ruland J. (2010) Recognition of RNA virus by RIG-I results in activation of CARD9 and inflammasome signaling for interleukin 1 beta production. *Nat Immunol* 11, 63-9.
Pohjoismaki J. L., Holmes J. B., Wood S. R., Yang M. Y., Yasukawa T., Reyes A., Bailey L. J., Cluett T. J., Goffart S., Willcox S., Rigby R. E., Jackson A. P., Spelbrink J. N., Griffith J. D., Crouch R. J., Jacobs H. T. and Holt I. J. (2010) Mammalian Mitochondrial DNA Replication Intermediates Are Essentially Duplex but Contain Extensive Tracts of RNA/DNA Hybrid. *J Mol Biol* 397, 1144-55.
Prichard M. N., Jairath S., Penfold M. E., St Jeor S., Bohlman M. C. and Pari G. S. (1998) Identification of persistent RNA-DNA hybrid structures within the origin of replication of human cytomegalovirus. *J Virol* 72, 6997-7004.
Prifti E., Zucker J. D., Clement K. and Henegar C. (2008) FunNet: an integrative tool for exploring transcriptional interactions. *Bioinformatics* 24, 2636-8.
Priyakumar U. D. and Mackerell A. D., Jr. (2008) Atomic detail investigation of the structure and dynamics of DNA.RNA hybrids: a molecular dynamics study. *J Phys Chem B* 112, 1515-24.
Prokunina L., Castillejo-Lopez C., Oberg F., Gunnarsson I., Berg L., Magnusson V., Brookes A. J., Tentler D., Kristjansdottir H., Grondal G., Bolstad A. I., Svenungsson E., Lundberg I., Sturfelt G., Jonssen A., Truedsson L., Lima G., Alcocer-Varela J., Jonsson R., Gyllensten U. B., Harley J. B., Alarcon-Segovia D., Steinsson K. and Alarcon-Riquelme M. E. (2002) A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 32, 666-9.
Pulendran B., Banchereau J., Burkeholder S., Kraus E., Guinet E., Chalouni C., Caron D., Maliszewski C., Davoust J., Fay J. and Palucka K. (2000) Flt3-ligand and granulocyte colony-stimulating factor mobilize distinct human dendritic cell subsets in vivo. *J Immunol* 165, 566-72.
Pursell Z. F., Isoz I., Lundstrom E. B., Johansson E. and Kunkel T. A. (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 317, 127-30.
Qiao F. and Bowie J. U. (2005) The many faces of SAM. *Sci STKE* 2005, re7.
Raines R. T. (1998) Ribonuclease A. *Chem Rev* 98, 1045-1066.

Ramantani G., Kohlhase J., Hertzberg C., Innes A. M., Engel K., Hunger S., Borozdin W., Mah J. K., Ungerath K., Walkenhorst H., Richardt H. H., Buckard J., Bevot A., Siegel C., von Stulpnagel C., Ikonomidou C., Thomas K., Proud V., Niemann F., Wieczorek D., Hausler M., Niggemann P., Baltaci V., Conrad K., Lebon P. and Lee-Kirsch M. A. (2010) Expanding the phenotypic spectrum of lupus erythematosus in Aicardi-Goutieres syndrome. *Arthritis Rheum* 62, 1469-77.

Rasmussen S. B., Jensen S. B., Nielsen C., Quartin E., Kato H., Chen Z. J., Silverman R. H., Akira S. and Paludan S. R. (2009) Herpes simplex virus infection is sensed by both Toll-like receptors and retinoic acid-inducible gene- like receptors, which synergize to induce type I interferon production. *J Gen Virol* 90, 74-8.

Ratmeyer L., Vinayak R., Zhong Y. Y., Zon G. and Wilson W. D. (1994) Sequence specific thermodynamic and structural properties for DNA.RNA duplexes. *Biochemistry* 33, 5298-304.

Raymond A. A., Zariah A. A., Samad S. A., Chin C. N. and Kong N. C. (1996) Brain calcification in patients with cerebral lupus. *Lupus* 5, 123-8.

Reaban M. E. and Griffin J. A. (1990) Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature* 348, 342-4.

Reaban M. E., Lebowitz J. and Griffin J. A. (1994) Transcription induces the formation of a stable RNA.DNA hybrid in the immunoglobulin alpha switch region. *J Biol Chem* 269, 21850-7.

Rebouillat D. and Hovanessian A. G. (1999) The human 2',5'-oligoadenylate synthetase family: interferon-induced proteins with unique enzymatic properties. *J Interferon Cytokine Res* 19, 295-308.

Rehli M. (2002) Of mice and men: species variations of Toll-like receptor expression. *Trends Immunol* 23, 375-8.

Rehwinkel J. and Reis e Sousa C. (2010) RIGorous detection: exposing virus through RNA sensing. *Science* 327, 284-6.

Rehwinkel J., Tan C. P., Goubau D., Schulz O., Pichlmair A., Bier K., Robb N., Vreede F., Barclay W., Fodor E. and Reis e Sousa C. (2010) RIG-I Detects Viral Genomic RNA during Negative-Strand RNA Virus Infection. *Cell* 140, 397 - 408.

Reis e Sousa C. (2004) Toll-like receptors and dendritic cells: for whom the bug tolls. *Semin Immunol* 16, 27-34.

Ren J. and Chaires J. B. (1999) Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry* 38, 16067-75.

Rice G., Newman W. G., Dean J., Patrick T., Parmar R., Flintoff K., Robins P., Harvey S., Hollis T., O'Hara A., Herrick A. L., Bowden A. P., Perrino F. W., Lindahl T., Barnes D. E. and Crow Y. J. (2007a) Heterozygous mutations in TREX1 cause familial chilblain lupus and dominant Aicardi-Goutieres syndrome. *Am J Hum Genet* 80, 811-5.

Rice G., Patrick T., Parmar R., Taylor C. F., Aeby A., Aicardi J., Artuch R., Montalto S. A., Bacino C. A., Barroso B., Baxter P., Benko W. S., Bergmann C., Bertini E., Biancheri R., Blair E. M., Blau N., Bonthron D. T., Briggs T., Brueton L. A., Brunner H. G., Burke C. J., Carr I. M., Carvalho D. R., Chandler K. E., Christen H. J., Corry P. C., Cowan F. M., Cox H., D'Arrigo S., Dean J., De Laet C., De Praeter C., Dery C., Ferrie C. D., Flintoff K., Frints S. G., Garcia-Cazorla A., Gener B., Goizet C., Goutieres F., Green A. J., Guet A., Hamel B. C., Hayward B. E., Heiberg A., Hennekam R. C., Husson M., Jackson A. P., Jayatunga R., Jiang Y. H., Kant S. G., Kao A., King M. D., Kingston H. M., Klepper J., van der Knaap M. S., Kornberg A. J., Kotzot D., Kratzer W., Lacombe D., Lagae L., Landrieu P. G., Lanzi G., Leitch A., Lim M. J., Livingston J. H., Lourenco C. M., Lyall E. G., Lynch S. A., Lyons M. J., Marom D., McClure J. P., McWilliam R., Melancon S. B., Mewasingh L. D., Moutard M. L., Nischal K. K., Ostergaard J. R., Prendiville J., Rasmussen M., Rogers R. C., Roland D., Rosser E. M., Rostasy K., Roubertie A., Sanchis A., Schiffmann R., Scholl-Burgi S., Seal S., Shalev S. A., Corcoles C. S., Sinha G. P., Soler D., Spiegel R., Stephenson J. B., Tacke U., Tan T. Y., Till M., Tolmie J. L., *et al.* (2007b) Clinical and molecular phenotype of Aicardi-Goutieres syndrome. *Am J Hum Genet* 81, 713-25.

Rice G. I., Bond J., Asipu A., Brunette R. L., Manfield I. W., Carr I. M., Fuller J. C., Jackson R. M., Lamb T., Briggs T. A., Ali M., Gornall H., Couthard L. R., Aeby A., Attard-Montalto S. P., Bertini E., Bodemer C., Brockmann K., Brueton L. A., Corry P. C., Desguerre I., Fazzi E., Cazorla A. G., Gener B., Hamel B. C., Heiberg A., Hunter M., van der Knaap M. S., Kumar R., Lagae L., Landrieu P. G., Lourenco C. M., Marom D., McDermott M. F., van der Merwe W., Orcesi S., Prendiville J. S., Rasmussen M., Shalev S. A., Soler D. M., Shinawi M., Spiegel R., Tan T. Y., Vanderver A., Wakeling E. L., Wassmer E., Whittaker E., Lebon P., Stetson D. B., Bonthron D. T. and Crow Y. J. (2009) Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat Genet* 41, 829-32.

Rice P. A. and Baker T. A. (2001) Comparative architecture of transposase and integrase complexes. *Nat Struct Biol* 8, 302-7.

Richards A., van den Maagdenberg A. M., Jen J. C., Kavanagh D., Bertram P., Spitzer D., Liszewski M. K., Barilla-Labarca M. L., Terwindt G. M., Kasai Y., McLellan M., Grand M. G., Vanmolkot K. R., de Vries B., Wan J., Kane M. J., Mamsa H., Schafer R., Stam A. H., Haan J., de Jong P. T., Storimans C. W., van Schooneveld M. J., Oosterhuis J. A., Gschwendter A., Dichgans M., Kotschet K. E., Hodgkinson S., Hardy T. A., Delatycki M. B., Hajj-Ali R. A., Kothari P. H., Nelson S. F., Frants R. R., Baloh R. W., Ferrari M. D. and Atkinson J. P. (2007) C-terminal truncations in human 3'-5' DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy. *Nat Genet* 39, 1068-1070.

Rigby R. E., Leitch A. and Jackson A. P. (2008) Nucleic acid-mediated inflammatory diseases. *Bioessays* 30, 833-42.

Roberts T. L., Idris A., Dunn J. A., Kelly G. M., Burnton C. M., Hodgson S., Hardy L. L., Garceau V., Sweet M. J., Ross I. L., Hume D. A. and Stacey K. J. (2009) HIN-200 Proteins Regulate Caspase Activation in Response to Foreign Cytoplasmic DNA. *Science* 323, 1057-60.

Roberts T. L., Sweet M. J., Hume D. A. and Stacey K. J. (2005) Cutting edge: species-specific TLR9-mediated recognition of CpG and non-CpG phosphorothioate-modified oligonucleotides. *J Immunol* 174, 605-8.

Rock F. L., Hardiman G., Timans J. C., Kastelein R. A. and Bazan J. F. (1998) A family of human receptors structurally related to Drosophila Toll. *Proc Natl Acad Sci U S A* 95, 588-93.

Rohman M. S., Koga Y., Takano K., Chon H., Crouch R. J. and Kanaya S. (2008) Effect of the disease-causing mutations identified in human ribonuclease (RNase) H2 on the activities and stabilities of yeast RNase H2 and archaeal RNase HII. *Febs J* 275, 4836-49.

Ronnblom L. and Alm G. V. (2003) Systemic lupus erythematosus and the type I interferon system. *Arthritis Res Ther* 5, 68-75.

Rosler A., Pohl M., Braune H. J., Oertel W. H., Gemsa D. and Sprenger H. (1998) Time course of chemokines in the cerebrospinal fluid and serum during herpes simplex type 1 encephalitis. *J Neurol Sci* 157, 82-9.

Rossi M. L. and Bambara R. A. (2006) Reconstituted Okazaki fragment processing indicates two pathways of primer removal. *J Biol Chem* 281, 26051-61.

Roth M. J., Tanese N. and Goff S. P. (1985) Purification and characterization of murine retroviral reverse transcriptase expressed in Escherichia coli. *J Biol Chem* 260, 9326-35.

Rothenfusser S., Goutagny N., DiPerna G., Gong M., Monks B. G., Schoenemeyer A., Yamamoto M., Akira S. and Fitzgerald K. A. (2005) The RNA helicase Lgp2 inhibits TLR-independent sensing of viral replication by retinoic acid-inducible gene-I. *J Immunol* 175, 5260-8.

Roy D., Zhang Z., Lu Z., Hsieh C. L. and Lieber M. R. (2010) Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site. *Mol Cell Biol* 30, 146-59.

Rubartelli A. and Lotze M. T. (2007) Inside, outside, upside down: damage-associated molecular-pattern molecules (DAMPs) and redox. *Trends Immunol* 28, 429-36.

Rydberg B. and Game J. (2002) Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc Natl Acad Sci U S A* 99, 16654-9.

Sabbah A., Chang T. H., Harnack R., Frohlich V., Tominaga K., Dube P. H., Xiang Y. and Bose S. (2009) Activation of innate immune antiviral responses by Nod2. *Nat Immunol* 10, 1073 - 1080.

Sadler A. J. and Williams B. R. (2008) Interferon-inducible antiviral effectors. *Nat Rev Immunol* 8, 559-68.

Saito T. and Gale M., Jr. (2008) Differential recognition of double-stranded RNA by RIG-I-like receptors in antiviral immunity. *J Exp Med* 205, 1523-7.
Saito T., Owen D. M., Jiang F., Marcotrigiano J. and Gale M. (2008) Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. *Nature* 454, 523-7.
Salazar M., Fedoroff O. Y., Miller J. M., Ribeiro N. S. and Reid B. R. (1993) The DNA strand in DNA.RNA hybrid duplexes is neither B-form nor A-form in solution. *Biochemistry* 32, 4207-15.
Samarajiwa S. A., Forster S., Auchettl K. and Hertzog P. J. (2009) INTERFEROME: the database of interferon regulated genes. *Nucleic Acids Res* 37, D852-7.
Sambrook J. and Russell D. W. (2000) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratories Press, US.
Sanghani S. R. and Lavery R. (1994) Theoretical studies of DNA-RNA hybrid conformations. *Nucleic Acids Res* 22, 1444-9.
Satoh T., Kato H., Kumagai Y., Yoneyama M., Sato S., Matsushita K., Tsujimura T., Fujita T., Akira S. and Takeuchi O. (2010) LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proc Natl Acad Sci U S A* 107, 1512-7.
Schenborn E. T. and Mierendorf R. C., Jr. (1985) A novel transcription property of SP6 and T7 RNA polymerases: dependence on template structure. *Nucleic Acids Res* 13, 6223-36.
Schlee M., Hartmann E., Coch C., Wimmenauer V., Janke M., Barchet W. and Hartmann G. (2009a) Approaching the RNA ligand for RIG-I? *Immunol Rev* 227, 66-74.
Schlee M. and Hartmann G. (2010) The chase for the RIG-I ligand--recent advances. *Mol Ther* 18, 1254-62.
Schlee M., Roth A., Hornung V., Hagmann C. A., Wimmenauer V., Barchet W., Coch C., Janke M., Mihailovic A., Wardle G., Juranek S., Kato H., Kawai T., Poeck H., Fitzgerald K. A., Takeuchi O., Akira S., Tuschl T., Latz E., Ludwig J. and Hartmann G. (2009b) Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* 31, 25-34.
Schmidt A., Schwerd T., Hamm W., Hellmuth J. C., Cui S., Wenzel M., Hoffmann F. S., Michallet M. C., Besch R., Hopfner K. P., Endres S. and Rothenfusser S. (2009) 5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I. *Proc Natl Acad Sci U S A* 106, 12067-72.
Schoeftner S. and Blasco M. A. (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* 10, 228-36.
Schulz O., Pichlmair A., Rehwinkel J., Rogers N. C., Scheuner D., Kato H., Takeuchi O., Akira S., Kaufman R. J. and Reis e Sousa C. (2010) Protein kinase R contributes to immunity against specific viruses by regulating interferon mRNA integrity. *Cell Host Microbe* 7, 354-61.
Sen G. C. and Sarkar S. N. (2005) Transcriptional signaling by double-stranded RNA: role of TLR3. *Cytokine Growth Factor Rev.* 16, 1-14.
Sester D. P., Naik S., Beasley S. J., Hume D. A. and Stacey K. J. (2000) Phosphorothioate backbone modification modulates macrophage activation by CpG DNA. *J Immunol* 165, 4165-73.
Seth R. B., Sun L., Ea C. K. and Chen Z. J. (2005) Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* 122, 669-82.
Shaban N. M., Harvey S., Perrino F. W. and Hollis T. (2010) The structure of the mammalian RNase H2 complex provides insight into RNA.DNA hybrid processing to prevent immune dysfunction. *J Biol Chem* 285, 3617-24.
Shaw N. N. and Arya D. P. (2008) Recognition of the unique structure of DNA:RNA hybrids. *Biochimie* 90, 1026-39.
Sheppard P., Kindsvogel W., Xu W., Henderson K., Schlutsmeyer S., Whitmore T. E., Kuestner R., Garrigues U., Birks C., Roraback J., Ostrander C., Dong D., Shin J., Presnell S., Fox B., Haldeman B., Cooper E., Taft D., Gilbert T., Grant F. J., Tackett M., Krivan W., McKnight G., Clegg C., Foster D. and Klucher K. M. (2003) IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat Immunol* 4, 63-8.
Shibata H., Tahira T. and Hayashi K. (1995) RNA-primed PCR. *Genome Res* 5, 400-3.
Shin H. D., Park B. L., Kim L. H., Lee H. S., Kim T. Y. and Bae S. C. (2004) Common DNase I polymorphism associated with autoantibody production among systemic lupus erythematosus patients. *Hum Mol Genet* 13, 2343-50.

Shin J. H. and Kelman Z. (2006) The replicative helicases of bacteria, archaea, and eukarya can unwind RNA-DNA hybrid substrates. *J Biol Chem* 281, 26914-21.
Shippen-Lentz D. and Blackburn E. H. (1990) Functional evidence for an RNA template in telomerase. *Science* 247, 546-52.
Sidorov A. V., Grasby J. A. and Williams D. M. (2004) Sequence-specific cleavage of RNA in the absence of divalent metal ions by a DNAzyme incorporating imidazolyl and amino functionalities. *Nucleic Acids Res* 32, 1591-601.
Siebke J. C. and Ekren T. (1970) Chromatography of RNA-DNA complexes on hydroxyapatite. A method for the separation of the complementary strands in T2 DNA. *Eur J Biochem* 12, 380-6.
Siegal F. P., Kadowaki N., Shodell M., Fitzgerald-Bocarsly P. A., Shah K., Ho S., Antonenko S. and Liu Y. J. (1999) The nature of the principal type 1 interferon-producing cells in human blood. *Science* 284, 1835-1837.
Siemer D., Kurth J., Lang S., Lehnerdt G., Stanelle J. and Kuppers R. (2008) EBV transformation overrides gene expression patterns of B cell differentiation stages. *Mol Immunol* 45, 3133-41.
Singer V. L., Jones L. J., Yue S. T. and Haugland R. P. (1997) Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. *Anal Biochem* 249, 228-38.
Singh R. R. (2005) SLE: translating lessons from model systems to human disease. *Trends Immunol* 26, 572-9.
Sioud M. (2006) Single-stranded small interfering RNA are more immunostimulatory than their double-stranded counterparts: a central role for 2'-hydroxyl uridines in immune responses. *Eur J Immunol* 36, 1222-30.
Smyth G. K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3, Article3.
Song J. J., Smith S. K., Hannon G. J. and Joshua-Tor L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* 305, 1434-7.
Sorensen T. L., Trebst C., Kivisakk P., Klaege K. L., Majmudar A., Ravid R., Lassmann H., Olsen D. B., Strieter R. M., Ransohoff R. M. and Sellebjerg F. (2002) Multiple sclerosis: a study of CXCL10 and CXCR3 co-localization in the inflamed central nervous system. *J Neuroimmunol* 127, 59-68.
Sorrentino S. (1998) Human extracellular ribonucleases: multiplicity, molecular diversity and catalytic properties of the major RNase types. *Cell Mol Life Sci* 54, 785-94.
Stein C. A., Subasinghe C., Shinozuka K. and Cohen J. S. (1988) Physicochemical properties of phosphorothioate oligodeoxynucleotides. *Nucleic Acids Res* 16, 3209-21.
Stetson D. B., Ko J. S., Heidmann T. and Medzhitov R. (2008) Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* 134, 587-98.
Stetson D. B. and Medzhitov R. (2006) Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* 24, 93-103.
Stith C. M., Sterling J., Resnick M. A., Gordenin D. A. and Burgers P. M. (2008) Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. *J Biol Chem* 283, 34129-40.
Storici F., Bebenek K., Kunkel T. A., Gordenin D. A. and Resnick M. A. (2007) RNA-templated DNA repair. *Nature* 447, 338-41.
Sugiyama T., Hoshino K., Saito M., Yano T., Sasaki I., Yamazaki C., Akira S. and Kaisho T. (2008) Immunoadjuvant effects of polyadenylic:polyuridylic acids through TLR3 and TLR7. *Int Immunol* 20, 1-9.
Summers K. and Crespi B. (2008) Molecular evolution of the prostate cancer susceptibility locus RNASEL: evidence for positive selection. *Infect Genet Evol* 8, 297-301.
Sun K., Coic E., Zhou Z., Durrens P. and Haber J. E. (2002) Saccharomyces forkhead protein Fkh1 regulates donor preference during mating-type switching through the recombination enhancer. *Genes Dev.* 16, 2085-2096.
Sutton A. and Freiman R. (1997) The Cak1p protein kinase is required at G1/S and G2/M in the budding yeast cell cycle. *Genetics* 147, 57-71.
Sutton D. H., Conn G. L., Brown T. and Lane A. N. (1997) The dependence of DNase I activity on the conformation of oligodeoxynucleotides. *Biochem J* 321 (Pt 2), 481-6.

Szekvolgyi L., Rakosy Z., Balint B. L., Kokai E., Imre L., Vereb G., Bacso Z., Goda K., Varga S., Balazs M., Dombradi V., Nagy L. and Szabo G. (2007) Ribonucleoprotein-masked nicks at 50-kbp intervals in the eukaryotic genomic DNA. *Proc Natl Acad Sci U S A* 104, 14964-9.

Tabeta K., Georgel P., Janssen E., Du X., Hoebe K., Crozat K., Mudd S., Shamel L., Sovath S., Goode J., Alexopoulou L., Flavell R. A. and Beutler B. (2004) Toll-like receptors 9 and 3 as essential components of innate immune defense against mouse cytomegalovirus infection. *Proc Natl Acad Sci U S A* 101, 3516-21.

Tabeta K., Hoebe K., Janssen E. M., Du X., Georgel P., Crozat K., Mudd S., Mann N., Sovath S., Goode J., Shamel L., Herskovits A. A., Portnoy D. A., Cooke M., Tarantino L. M., Wiltshire T., Steinberg B. E., Grinstein S. and Beutler B. (2006) The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9. *Nat Immunol* 7, 156-64.

Tadokoro T. and Kanaya S. (2009) Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *FEBS J* 276, 1482-93.

Takahashi K., Shibata T., Akashi-Takamura S., Kiyokawa T., Wakabayashi Y., Tanimura N., Kobayashi T., Matsumoto F., Fukui R., Kouro T., Nagai Y., Takatsu K., Saitoh S. and Miyake K. (2007) A protein associated with Toll-like receptor (TLR) 4 (PRAT4A) is required for TLR-dependent immune responses. *J Exp Med* 204, 2963-76.

Takahasi K., Yoneyama M., Nishihori T., Hirai R., Kumeta H., Narita R., Gale M., Jr., Inagaki F. and Fujita T. (2008) Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses. *Mol Cell* 29, 428-40.

Takaoka A. and Taniguchi T. (2008) Cytosolic DNA recognition for triggering innate immune responses. *Adv Drug Deliv Rev* 60, 847-57.

Takaoka A., Wang Z., Choi M. K., Yanai H., Negishi H., Ban T., Lu Y., Miyagishi M., Kodama T., Honda K., Ohba Y. and Taniguchi T. (2007) DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* 448, 501-5.

Takaoka A., Yanai H., Kondo S., Duncan G., Negishi H., Mizutani T., Kano S., Honda K., Ohba Y., Mak T. W. and Taniguchi T. (2005) Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* 434, 243-9.

Talal N. and Gallo R. C. (1972) Antibodies to a DNA:RNA hybrid in systemic lupus erythematosus measured by a cellulose ester filter radioimmunoassay. *Nat New Biol* 240, 240-2.

Tanese N., Telesnitsky A. and Goff S. P. (1991) Abortive reverse transcription by mutants of Moloney murine leukemia virus deficient in the reverse transcriptase-associated RNase H function. *J Virol* 65, 4387-97.

Taniguchi T. and Takaoka A. (2002) The interferon-alpha/beta system in antiviral responses: a multimodal machinery of gene regulation by the IRF family of transcription factors. *Curr Opin Immunol* 14, 111-6.

Tataurov A. V., You Y. and Owczarzy R. (2008) Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys Chem* 133, 66-70.

Thomas M., White R. L. and Davis R. W. (1976) Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc Natl Acad Sci U S A* 73, 2294-8.

Tian J., Avalos A. M., Mao S. Y., Chen B., Senthil K., Wu H., Parroche P., Drabic S., Golenbock D., Sirois C., Hua J., An L. L., Audoly L., La Rosa G., Bierhaus A., Naworth P., Marshak-Rothstein A., Crow M. K., Fitzgerald K. A., Latz E., Kiener P. A. and Coyle A. J. (2007) Toll-like receptor 9-dependent activation by DNA-containing immune complexes is mediated by HMGB1 and RAGE. *Nat Immunol* 8, 487-96.

Ting J. P., Duncan J. A. and Lei Y. (2010) How the noninflammasome NLRs function in the innate immune system. *Science* 327, 286-90.

Tluk S., Jurk M., Forsbach A., Weeratna R., Samulowitz U., Krieg A. M., Bauer S. and Vollmer J. (2009) Sequences derived from self-RNA containing certain natural modifications act as suppressors of RNA-mediated inflammatory immune responses. *Int Immunol* 21, 607-19.

Tous C. and Aguilera A. (2007) Impairment of transcription elongation by R-loops in vitro. *Biochem Biophys Res Commun* 360, 428-32.

Traganos F., Darzynkiewicz Z., Sharpless T. and Melamed M. R. (1977) Simultaneous staining of ribonucleic and deoxyribonucleic acids in unfixed cells using acridine orange in a flow cytofluorometric system. *J Histochem Cytochem* 25, 46-56.

Triana-Alonso F. J., Dabrowski M., Wadzack J. and Nierhaus K. H. (1995) Self-coded 3'-extension of run-off transcripts produces aberrant products during in vitro transcription with T7 RNA polymerase. *J Biol Chem* 270, 6298-307.
Trifilo M. J. and Lane T. E. (2003) Adenovirus-mediated expression of CXCL10 in the central nervous system results in T-cell recruitment and limited neuropathology. *J Neurovirol* 9, 315-24.
Tsurimoto T. and Stillman B. (1991) Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis. *J Biol Chem* 266, 1961-8.
Tuma R. S., Beaudet M. P., Jin X., Jones L. J., Cheung C. Y., Yue S. and Singer V. L. (1999) Characterization of SYBR Gold nucleic acid gel stain: a dye optimized for use with 300-nm ultraviolet transilluminators. *Anal Biochem* 268, 278-88.
Turchi J. J., Huang L., Murante R. S., Kim Y. and Bambara R. A. (1994) Enzymatic completion of mammalian lagging-strand DNA replication. *Proc.Natl.Acad.Sci.U.S.A* 91, 9803-9807.
Untergasser A., Nijveen H., Rao X., Bisseling T., Geurts R. and Leunissen J. A. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* 35, W71-4.
Unterholzner L., Keating S. E., Baran M., Horan K. A., Jensen S. B., Sharma S., Sirois C. M., Jin T., Latz E., Xiao T. S., Fitzgerald K. A., Paludan S. R. and Bowie A. G. (2010) IFI16 is an innate immune sensor for intracellular DNA. *Nat Immunol* 11, 997-1004.
Van C., Yan S., Michael W. M., Waga S. and Cimprich K. A. (2010) Continued primer synthesis at stalled replication forks contributes to checkpoint activation. *J Cell Biol* 189, 233-46.
van Heteren J. T., Rozenberg F., Aronica E., Troost D., Lebon P. and Kuijpers T. W. (2008) Astrocytes produce interferon-alpha and CXCL10, but not IL-6 or CXCL8, in aicardi-Goutieres syndrome. *Glia* 56, 568-78.
Van Heteren J. T., van der Knaap M. S., Poll The B. W. and Kuijpers T. W. (2007) Plasmacytoid dendritic cells and interferon-alpha in Aicardi-Goutieres syndrome. *Neuropediatrics* 38, 269-275.
van Pesch V., Lanaya H., Renauld J. C. and Michiels T. (2004) Characterization of the murine alpha interferon gene family. *J Virol* 78, 8219-28.
Vengrova S. and Dalgaard J. Z. (2004) RNase-sensitive DNA modification(s) initiates S. pombe mating-type switching. *Genes Dev* 18, 794-804.
Vengrova S. and Dalgaard J. Z. (2006) The wild-type Schizosaccharomyces pombe mat1 imprint consists of two ribonucleotides. *EMBO Rep* 7, 59-65.
Verrips A., Hiel J. A., Gabreels F. J., Wesseling P. and Rotteveel J. J. (1997) The Aircardi-Goutieres syndrome: variable clinical expression in two siblings. *Pediatr.Neurol.* 16, 323-325.
Vilaysane A. and Muruve D. A. (2009) The innate immune response to DNA. *Semin Immunol* 21, 208 - 214.
Vogt V. M. (1973) Purification and further properties of single-strand-specific nuclease from Aspergillus oryzae. *Eur J Biochem* 33, 192-200.
Vollmer J., Tluk S., Schmitz C., Hamm S., Jurk M., Forsbach A., Akira S., Kelly K. M., Reeves W. H., Bauer S. and Krieg A. M. (2005) Immune stimulation mediated by autoantigen binding sites within small nuclear RNAs involves Toll-like receptors 7 and 8. *J Exp Med* 202, 1575-85.
Vyse T. J. and Kotzin B. L. (1998) Genetic susceptibility to systemic lupus erythematosus. *Annu Rev Immunol* 16, 261-92.
Waga S. and Stillman B. (1994) Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. *Nature* 369, 207-12.
Waga S. and Stillman B. (1998) The DNA replication fork in eukaryotic cells. *Annu Rev Biochem* 67, 721-51.
Wang A. H., Fujii S., van Boom J. H., van der Marel G. A., van Boeckel S. A. and Rich A. (1982) Molecular structure of r(GCG)d(TATACGC): a DNA--RNA hybrid helix joined to double helical DNA. *Nature* 299, 601-4.
Wang T., Town T., Alexopoulou L., Anderson J. F., Fikrig E. and Flavell R. A. (2004) Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nat Med* 10, 1366-73.

Wang Y., Ludwig J., Schuberth C., Goldeck M., Schlee M., Li H., Juranek S., Sheng G., Micura R., Tuschl T., Hartmann G. and Patel D. J. (2010a) Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat Struct Mol Biol* 17, 781-7.

Wang Y. Y., Liu L. J., Zhong B., Liu T. T., Li Y., Yang Y., Ran Y., Li S., Tien P. and Shu H. B. (2010b) WDR5 is essential for assembly of the VISA-associated signaling complex and virus-triggered IRF3 and NF-kappaB activation. *Proc Natl Acad Sci U S A* 107, 815-20.

Wassmer E., Singh J., Agrawal S., Santra S. and Crow Y. J. (2009) Elevated CSF pterins - biochemical marker of Aicardi-Goutieres syndrome. *Dev Med Child Neurol* 51, 841-2.

Wathelet M. G., Berr P. M. and Huez G. A. (1992) Regulation of gene expression by cytokines and virus in human cells lacking the type-I interferon locus. *Eur J Biochem* 206, 901-10.

Wheelhouse R. T. and Chaires J. B. (2010) Drug binding to DNA x RNA hybrid structures. *Methods Mol Biol* 613, 55-70.

Wilson F. E. and Sugino A. (1985) Purification of a DNA primase activity from the yeast Saccharomyces cerevisiae. Primase can be separated from DNA polymerase I. *J Biol Chem* 260, 8173-81.

Wohrl B. M. and Molling K. (1990) Interaction of HIV-1 ribonuclease H with polypurine tract containing RNA-DNA hybrids. *Biochemistry* 29, 10141-7.

Wotton D. and Shore D. (1997) A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in Saccharomyces cerevisiae. *Genes Dev* 11, 748-60.

Wu H., Lima W. F. and Crooke S. T. (2001) Investigating the structure of human RNase H1 by site-directed mutagenesis. *J Biol Chem* 276, 23547-53.

Wu H., Xu H., Miraglia L. J. and Crooke S. T. (2000) Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. *J Biol Chem* 275, 36957-65.

Xiang Y., Wang Z., Murakami J., Plummer S., Klein E. A., Carpten J. D., Trent J. M., Isaacs W. B., Casey G. and Silverman R. H. (2003) Effects of RNase L mutations associated with prostate cancer on apoptosis induced by 2',5'-oligoadenylates. *Cancer Res* 63, 6795-801.

Xu B. and Clayton D. A. (1995) A persistent RNA-DNA hybrid is formed during transcription at a phylogenetically conserved mitochondrial DNA sequence. *Mol Cell Biol* 15, 580-9.

Xu B. and Clayton D. A. (1996) RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *Embo J* 15, 3135-43.

Xu L. G., Wang Y. Y., Han K. J., Li L. Y., Zhai Z. and Shu H. B. (2005) VISA is an adapter protein required for virus-triggered IFN-beta signaling. *Mol Cell* 19, 727-40.

Xu Y., Suzuki Y., Kaminaga K. and Komiyama M. (2009) Molecular basis of human telomere DNA/RNA structure and its potential application. *Nucleic Acids Symp Ser (Oxf)*, 63-4.

Yan N., Cherepanov P., Daigle J. E., Engelman A. and Lieberman J. (2009) The SET complex acts as a barrier to autointegration of HIV-1. *PLoS Pathog* 5, e1000327.

Yan N., Regalado-Magdos A. D., Stiggelbout B., Lee-Kirsch M. A. and Lieberman J. (2010) The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat Immunol* 11, 1005-13.

Yang D., Buchholz F., Huang Z., Goga A., Chen C. Y., Brodsky F. M. and Bishop J. M. (2002a) Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells. *Proc Natl Acad Sci U S A* 99, 9942-7.

Yang M. Y., Bowmaker M., Reyes A., Vergani L., Angeli P., Gringeri E., Jacobs H. T. and Holt I. J. (2002b) Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell* 111, 495-505.

Yang Y. G., Lindahl T. and Barnes D. E. (2007) Trex1 Exonuclease Degrades ssDNA to Prevent Chronic Checkpoint Activation and Autoimmune Disease. *Cell* 131, 873-86.

Yasuda K., Ogawa Y., Yamane I., Nishikawa M. and Takakura Y. (2005a) Macrophage activation by a DNA/cationic liposome complex requires endosomal acidification and TLR9-dependent and -independent pathways. *J Leukoc Biol* 77, 71-9.

Yasuda K., Richez C., Maciaszek J. W., Agrawal N., Akira S., Marshak-Rothstein A. and Rifkin I. R. (2007) Murine dendritic cell type I IFN production induced by human IgG-RNA immune complexes is IFN regulatory factor (IRF)5 and IRF7 dependent and is required for IL-6 production. *J Immunol* 178, 6876-85.

Yasuda K., Yu P., Kirschning C. J., Schlatter B., Schmitz F., Heit A., Bauer S., Hochrein H. and Wagner H. (2005b) Endosomal translocation of vertebrate DNA activates dendritic cells via TLR9-dependent and -independent pathways. *J Immunol* 174, 6129-36.
Yasuda T., Nadano D., Sawazaki K. and Kishi K. (1992) Genetic polymorphism of human deoxyribonuclease II (DNase II): low activity levels in urine and leukocytes are due to an autosomal recessive allele. *Ann Hum Genet* 56, 1-10.
Yasukawa T., Reyes A., Cluett T. J., Yang M. Y., Bowmaker M., Jacobs H. T. and Holt I. J. (2006) Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J* 25, 5358-71.
Yasutomo K., Horiuchi T., Kagami S., Tsukamoto H., Hashimura C., Urushihara M. and Kuroda Y. (2001) Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat Genet* 28, 313-4.
Yeh T. M., Chang H. C., Liang C. C., Wu J. J. and Liu M. F. (2003) Deoxyribonuclease-inhibitory antibodies in systemic lupus erythematosus. *J Biomed Sci* 10, 544-51.
Yoneyama M. and Fujita T. (2010) Recognition of viral nucleic acids in innate immunity. *Rev Med Virol* 20, 4-22.
Yoneyama M., Kikuchi M., Natsukawa T., Shinobu N., Imaizumi T., Miyagishi M., Taira K., Akira S. and Fujita T. (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* 5, 730-7.
Yoshida H., Okabe Y., Kawane K., Fukuyama H. and Nagata S. (2005) Lethal anemia caused by interferon-beta produced in mouse embryos carrying undigested DNA. *Nat Immunol* 6, 49-56.
Yoshinaga T., Yasuda K., Ogawa Y., Nishikawa M. and Takakura Y. (2007) DNA and its cationic lipid complexes induce CpG motif-dependent activation of murine dendritic cells. *Immunology* 120, 295-302.
Yu K., Chedin F., Hsieh C. L., Wilson T. E. and Lieber M. R. (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* 4, 442-51.
Yu K. and Lieber M. R. (2003) Nucleic acid structures and enzymes in the immunoglobulin class switch recombination mechanism. *DNA Repair (Amst)* 2, 1163-74.
Yuzhakov A., Kelman Z. and O'Donnell M. (1999) Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell* 96, 153-63.
Zeng W. and Chen Z. J. (2008) MITAgating viral infection. *Immunity* 29, 513-5.
Zhang Q., Raoof M., Chen Y., Sumi Y., Sursal T., Junger W., Brohi K., Itagaki K. and Hauser C. J. (2010) Circulating mitochondrial DAMPs cause inflammatory responses to injury. *Nature* 464, 104-7.
Zhang S. Y., Jouanguy E., Ugolini S., Smahi A., Elain G., Romero P., Segal D., Sancho-Shimizu V., Lorenzo L., Puel A., Picard C., Chapgier A., Plancoulaine S., Titeux M., Cognet C., von Bernuth H., Ku C. L., Casrouge A., Zhang X. X., Barreiro L., Leonard J., Hamilton C., Lebon P., Heron B., Vallee L., Quintana-Murci L., Hovnanian A., Rozenberg F., Vivier E., Geissmann F., Tardieu M., Abel L. and Casanova J. L. (2007) TLR3 deficiency in patients with herpes simplex encephalitis. *Science* 317, 1522-7.
Zheng L., Dai H., Zhou M., Li M., Singh P., Qiu J., Tsark W., Huang Q., Kernstine K., Zhang X., Lin D. and Shen B. (2007) Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nat Med* 13, 812-9.
Zheng L., Zhou M., Chai Q., Parrish J., Xue D., Patrick S. M., Turchi J. J., Yannone S. M., Chen D. and Shen B. (2005) Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. *EMBO Rep* 6, 83-9.
Zheng X. and Bevilacqua P. C. (2004) Activation of the protein kinase PKR by short double-stranded RNAs with single-stranded tails. *RNA* 10, 1934-45.
Zhong B., Yang Y., Li S., Wang Y. Y., Li Y., Diao F., Lei C., He X., Zhang L., Tien P. and Shu H. B. (2008) The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity* 29, 538-50.
Zhong B., Zhang L., Lei C., Li Y., Mao A. P., Yang Y., Wang Y. Y., Zhang X. L. and Shu H. B. (2009) The ubiquitin ligase RNF5 regulates antiviral responses by mediating degradation of the adaptor protein MITA. *Immunity* 30, 397-407.
Zhou A., Hassel B. A. and Silverman R. H. (1993) Expression cloning of 2-5A-dependent RNAase: a uniquely regulated mediator of interferon action. *Cell* 72, 753-65.

Zhou A., Paranjape J., Brown T. L., Nie H., Naik S., Dong B., Chang A., Trapp B., Fairchild R., Colmenares C. and Silverman R. H. (1997) Interferon action and apoptosis are defective in mice devoid of 2',5'-oligoadenylate-dependent RNase L. *Embo J* 16, 6355-63.

Zhuang H., Kosboth M., Lee P., Rice A., Driscoll D. J., Zori R., Narain S., Lyons R., Satoh M., Sobel E. and Reeves W. H. (2006) Lupus-like disease and high interferon levels corresponding to trisomy of the type I interferon cluster on chromosome 9p. *Arthritis Rheum* 54, 1573-9.

Zimmerman S. B. and Pheiffer B. H. (1981) A RNA.DNA hybrid that can adopt two conformations: an x-ray diffraction study of poly(rA).poly(dT) in concentrated solution or in fibers. *Proc Natl Acad Sci U S A* 78, 78-82.

Zipper H., Brunner H., Bernhagen J. and Vitzthum F. (2004) Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res* 32, e103.

Zou H. Y., Ma L., Meng M. J., Yao X. S., Lin Y., Wu Z. Q., He X. W., Wang J. F. and Wang X. N. (2007) Expression of recombination-activating genes and T cell receptor gene recombination in the human T cell leukemia cell line. *Chin Med J (Engl)* 120, 410-5.

Zuker M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31, 3406-15.